

ISSN (Print): 2077-7973  
ISSN (Online): 2077-8767  
DOI: 10.6977/IJoSI.202502\_9(1)

# International Journal of Systematic Innovation



VOL. 09 NO. 01  
February, 2025

Published by the Society of Systematic Innovation

---

***Opportunity Identification  
&  
Problem Solving***

# The International Journal of Systematic Innovation

---

**Publisher:**

The Society of Systematic Innovation

**Editorial Team:**Editor-in-Chief:

Sheu, Dongliang Daniel (National Tsing Hua University, Taiwan)

Executive Editor:

Deng, Jyhjeng (DaYeh University, Taiwan)

Yeh, W. C. (National Tsing Hua University, Taiwan)

Editorial Team Members (in alphabetical order):

- Cavallucci, Denis (INSA Strasbourg University, France)
- De Guio, Roland (INSA Strasbourg University, France)
- Feygenson, Oleg (Algorithm Technology Research Center, Russian Federation)
- Filmore, Paul (University of Plymouth, UK)
- Kusiak, Andrew (University of Iowa, USA)
- Lee, Jay (University of Cincinnati, USA)
- Litvin, Simon (GEN TRIZ, USA)
- Lu, Stephen (University of Southern California, USA)
- Mann, Darrell (Ideal Final Result, Inc., UK)

- Sawaguch, Manabu (Waseda University, Japan)
- Shouchkov, Valeri (ICG Training& Consulting, Netherlands)
- Song, Yong-Won (Korea Polytechnic University, Korea)
- Yu, Oliver (San Jose State University, USA)
- Zhang, Zhinan (Shanghai Jiao Tong University)

Managing Editor:

- Jimmy Li

Assistant Editors:

- Chiaoling Ni

**Editorial Office:**

- The International Journal of Systematic Innovation
- 6F, # 352, Sec. 2, Guanfu Rd, Hsinchu, Taiwan, R.O.C., 30071
- e-mail: editor@systematic-innovation.org  
ijosi@systematic-innovation.org
- web site: <http://www.IJoSI.org>
- Tel: +886-3572-3200

# INTERNATIONAL JOURNAL OF SYSTEMATIC INNOVATION

## CONTENTS

FEBRUARY 2025 VOLUME 9 ISSUE 1

---

### FULL PAPERS

- Evaluation of convolutional neural network models' performance for estimating mango crop yield  
..... M.V. Neethi, P. Raviraj. 1–18
- Technology innovation of dryer machine based on sustainability automation systems to increase agel fiber production in handicraft SME  
..... K. Maruf, R.J. Setiawan, Darmono, S.F.A. Widodo, S.S. Nugroho, N.E. Khosyati, N. Azizah. 19–29
- Brightness augmentation implementation to evaluate performance classification of masked facial expressions based on the CNN model  
..... D.M. Ramadhan, H. Mubarak, Rianto. 30–43
- A systematic meta-analysis on the role of artificial intelligence and machine learning in detection of gynecological disorders  
..... J.I. Nandalwar, P.M. Jaeandhiya. 44–56
- Systematic modernization of fish smoking method with the implementation of smoked fish machine based on Internet of Things technology  
..... R.J. Setiawan, K. Ma'ruf, Darmono, N. Azizah, N.E. Khosyati. 57–67
- Knowledge management capability and innovation ambidexterity: The role of intellectual capital and intangible competitive advantage  
..... N. Sirivariskul. 68–77
- Custom hardware design for peripheral artery disease detection: Field-programmable gate arrays and application-specific integrated circuits  
..... N. Pravalika, A. Jabeena, V. Rajamani. 78–97
- Predictive analytics: Unveiling the potential of machine learning and deep learning  
..... P. Kavitha, L. Shakkeera. 98–110
- MNETGIDD: A heuristic-oriented segmentation and deep learning multi-disease detection model for gastrointestinal tracts  
..... A. Bamini. 111–131

# Evaluation of convolutional neural network models' performance for estimating mango crop yield

M.V. Neethi<sup>1</sup>, P. Raviraj<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering-Data Science, ATME College of Engineering,  
Affiliated to Visvesvaraya Technological University (VTU), Belagavi, India

<sup>2</sup>Department of Computer Science and Engineering, GSSS Institute of Engineering and Technology for Women,  
Affiliated to Visvesvaraya Technological University (VTU), Belagavi, India

\* Corresponding author E-mail: neethi.mv@gmail.com

(Received 16 April 2024; Final version received 3 August 2024; Accepted 16 August 2024)

## Abstract

In agriculture, crop yield estimation is essential; producers, industrialists, and consumers all benefit from knowing the early yield. Manual mango counting typically involves the utilization of human labor. Experts visually examine each sample to complete the process, which is time-consuming, difficult, and lacks precision. For commercial mango production to produce high-quality fruits from the orchard to the consumer, a quick, non-destructive, and accurate variety classification is required. Because of its effectiveness in computer vision, a convolutional neural network—one of the deep learning techniques—was chosen for this investigation. For yield prediction, a total of eight popular mango cultivars were utilized. A comparison with previously trained models was used to assess the proposed model. The performance of the classifiers was evaluated using evaluation metrics such as accuracy, loss, area under the receiver operating characteristic curve score, precision, recall, F1-score, sensitivity, specificity, positive predictive value, negative predictive value, and Cohen's Kappa performance measure. In terms of performance evaluation criteria, it was found that the proposed approach outperformed the pre-trained models. The suggested model achieved 98.85% accuracy in the test set, which had 800 images. This outcome demonstrates the tangible applicability of the proposed methodology for mango crop estimation.

**Keywords:** Convolutional Neural Networks, Deep Learning, Early Detection, Machine Learning

## 1. Introduction

Precise prediction of crop production levels is essential for developing efficient farming practices and preserving food security in a rapidly changing world. Crop yield affects food productivity and is crucial in ensuring food availability and safety, as recognized by policymakers, farmers, and consumers

alike. Therefore, predicting crop yields offers a significant advantage in supporting financial and managerial decisions. Crop yield is the ratio of agricultural input to output, measured as the quantity of yield per unit area of cultivated land. Hence, crop yield is a typical indicator of agricultural productivity. Crop productivity can be affected by various factors,

such as pests and diseases, environmental changes, farming techniques, and consumer demand.

Computer vision-based automatic crop yield estimation greatly helps agronomists to identify and categorise plant diseases, allowing them to take necessary measures at an early stage to prevent further damage to the plant (Xin Zhang et al., 2019). A simple plant disease detection system can be used as an efficient tool by small-scale farmers for early disease detection. In large-scale cultivation, the system can be mounted on autonomous vehicles that cover the entire region with continuous video coverage or image capturing which is monitored timely and accurately by the agronomist. This approach can help detect, diagnose, and treat pathological problems to a great extent. Over the decades, many types of research have been developed in the field of plant disease detection. Numerous machine learning algorithms have been created, which learn iterative from the given data and build the model without explicit programming, providing an efficient tool for plant disease detection, not only for identification.

In this study, the proposed model was compared and analyzed with nine different convolutional pre-trained neural networks, namely: AlexNet, VGG16, VGG19, DenseNet-121, ResNet-101, MobileNetV2, Xception, NasNet-Mobile, and Inception-v networks. The experiments involved tweaking the hyperparameters to improve the performance of the convolutional neural network (CNN) model. The study hypothesized that deep learning models would help phytopathologist and farmers estimate yield at an early stage.

The contents of the paper are organized as follows: Section 2 provides a survey of the related literature. Section 3 presents the dataset for experimentation and testing, as well as experiments designed to study the elements that impact the overall performance and effectiveness of the evolutionary system. Section 4 presents the results and Section 5 provides an analysis of the proposed method. The article ends with the conclusion and recommendations.

## 2. Literature Survey

The literature contains various works related to crop yield estimation. The development of agricultural deep learning technology, especially in yield estimation detection, has only started in recent years and remains somewhat limited. Deep learning, a subset of machine learning, helps identify and classify different plant diseases (Xin Yang et al., 2016). Deep learning utilizes multi-level depiction and

generalization by employing multi-layer nonlinear modules for feature extraction and transformation, distinguishing it from traditional machine learning (Fine, 2006; Fuentes et al., 2017). CNN is a deep learning algorithm that takes an image as the input, learns the objects and patterns within the image, and differentiates it from others (Grinblat et al., 2016). A smartphone-assisted CNN model is used for crop disease detection using a publicly available dataset to identify 14 different crops and 26 different diseases, as CNN requires less pre-processing compared to other algorithms, concerning hyper-spectral images. However, CNN faces challenges in processing the high-dimensional information in multidimensional data cubes, which leads to high computational time (Mohanty et al., 2016; Paoletti et al., 2018). Nuclear discriminant analysis based on the Spectral Vegetation Index method is used for the detection and classification of yellow rust, aphids, and powdery mildew in winter wheat (Shi et al., 2017). In hyper-spectral reflecting datasets at the leaf and canopy levels, the model outperforms the traditional linear discriminant method for classifying healthy leaf and diseased wheat leaves. A leaf-based CNN model for disease recognition and classification successfully classified 13 different diseases, with precision ranging from 91% to 98% (Sladojevic et al., 2016). A dataset containing 79,265 images was introduced by Arsenovic et al., 2019). Two types of augmentation techniques were applied to increase the dataset size, and a new two-stage neural network architecture was proposed for the classification of plant diseases based on the current environment. The trained model achieved an accuracy of 93.67%. A method was developed to classify citrus disease using the  $\Delta E$  colour difference algorithm to segment the diseased area. This method achieved 99.9% accuracy (Ali et al., 2017). A novel cucumber disease recognition model was developed using three pipelined procedures: segmenting diseased leaves with K-means clustering, color and shape detection, and classification of diseased leaves by sparse representation. The developed method was compared with other methods and achieved a classification performance of 85.7% (Zhang et al., 2017). A new CNN architecture with two deep classifiers and a trainable visualization method for plant disease classification was proposed. The architecture trained two classifiers in parallel, and the area over perturbation curve was used to compare the proposed method with the existing state of art method, achieving a performance of approximately 0.907 (Brahimi et al., 2019). Different plant species were classified, and a content-based data retrieval method

was used to search for plant species using a deep learning approach (Gyires-Tóth et al., 2019). The EfficientNet deep learning architecture was proposed to detect plant diseases. Plant village datasets were tested on EfficientNet, achieving accuracies of 99.91% and 99.97% for the original images, and 98.42% and 99.39% for the enhanced images (Atila et al., 2021).

### 3. Materials and Methods

#### 3.1. Dataset

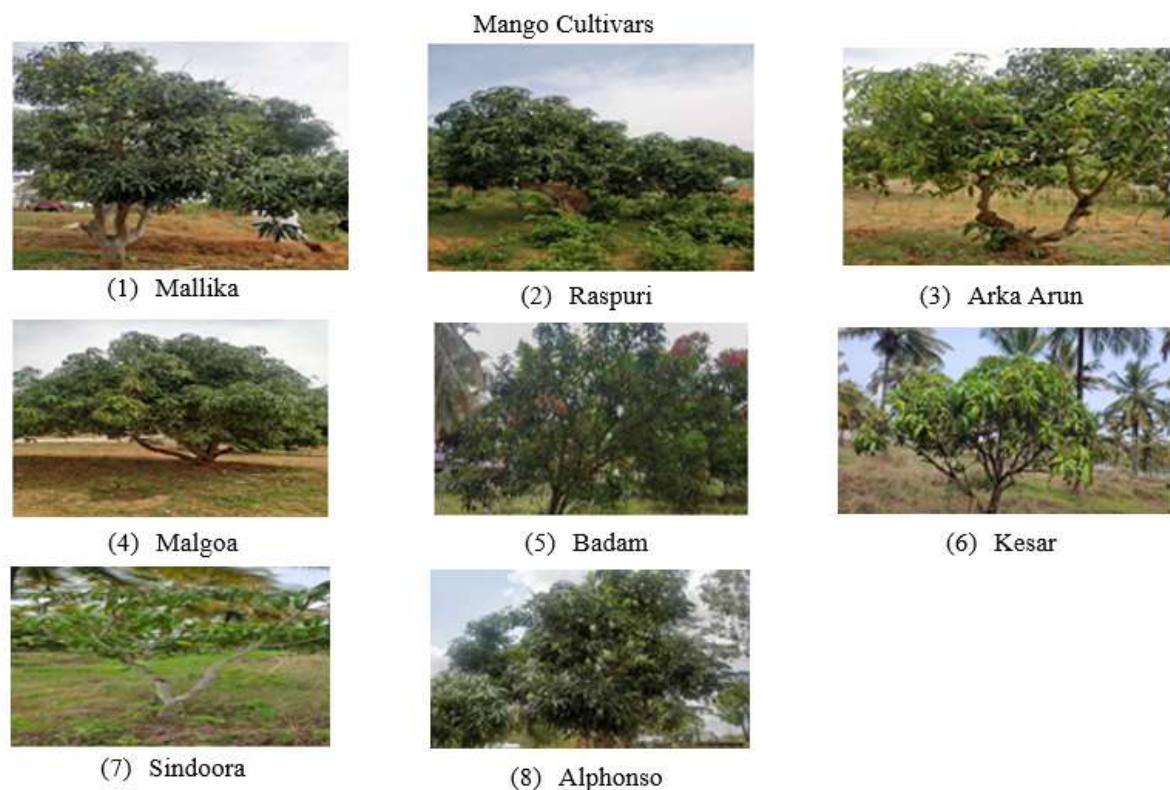
In this study, a temporal mango crop dataset of 4,000 images, including eight different cultivars, was used. The dataset of images, captured under real cultivation conditions, contains various objects like ground, other parts of the plant, varying illumination conditions, occlusions, etc. Fig. 1 and Table 1 depict the dataset and provide detailed information about it. A large dataset is required for the robust performance of CNN models. The mango fruit dataset was built under a limited set of conditions; to improve variability and increase the dataset size. Image augmentation techniques such as rotation, reflection, translation, scaling, and shear were applied to the dataset images for both the x- and y-axes. The entire dataset was initially divided into two subsets: the training set and the testing set, with images randomly

split in an 80/20 ratio. A Python script was developed to automatically divide dataset images into two sets, which includes size reduction and normalizing to a 224×224 pixel size. Fig. 2 provides an overview of the proposed workflow.

#### 3.2. Deep Learning Models

A CNN is a deep learning algorithm that takes an input image and assigns importance to various features and objects within the image to distinguish them. The pre-processing required for a convolution network is much lower than for other classification algorithms. Several innovative ideas have contributed to the evolution of CNNs, including optimization of parameters using various activation and loss functions, the development of standardized architectures, etc. (Khan et al., 2020; Zhou, 2020). An overview of the deep learning model is depicted in Fig. 3.

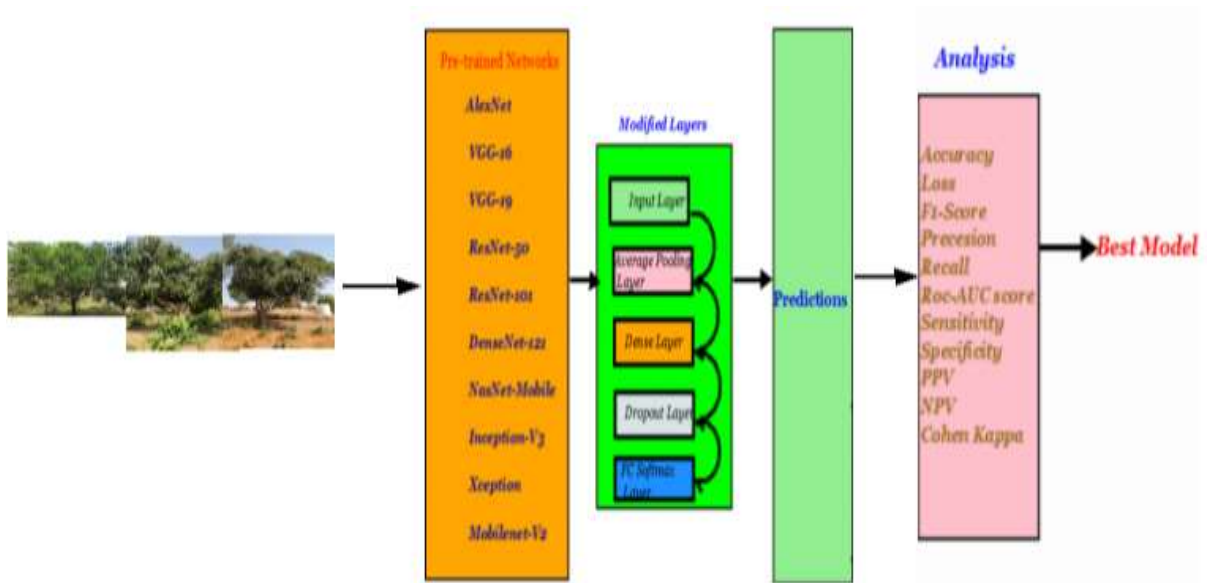
In this study, an FRCNN ResNet-50 neural network models were considered to estimate eight different cultivars of mango fruit crops and were compared with nine well-known pre-trained deep networks: AlexNet, VGG-16, VGG-19, ResNet-101, DenseNet-121, NasNet-Mobile, Inception-V3, Xception, and Mobilenet-V2.



**Fig. 1.** Dataset description.

**Table 1.** Detailed information of the dataset.

SL. no.	Image dataset	Number of samples
1	Mallika	500
2	Raspuri	500
3	Arka Arun	500
4	Malgoa	500
5	Badam	500
6	Kesar	500
7	Sindoorra	500
8	Alphanso	500



**Fig. 2.** Flowchart of the overall study. Abbreviations: NPV: Negative predictive value; PPV: Positive predictive value; Roc-AUC: Area under the receiver operating characteristic curve.



AlexNet, designed by Alex Krizhevsky, is a feedforward CNN architecture consisting of eight layers: five convolutional layers containing several kernels and three fully connected layers (f6, f7 and f8) (Krizhevsky et al., 2012). VGG-16 is a deep CNN with a depth of 16 layers, consisting of five convolutional blocks (13 layers) and three fully connected layers. The network consists of five larger blocks, proposed by Karen Simonyan (Simonyan et al. 2015). VGG19 is deeper than VGG 16, with 19 layers in total. It has five convolutional blocks (16 layers) followed by three fully connected layers. The ResNet CNN model is made up of residual blocks. ResNet-50 is a CNN with 50 layers, consisting of 48 convolutional layers, one MaxPool layer, and one average pool layer. The depth of ResNet-101 is 101 layers (He et al., 2016). DenseNet-121, also a CNN model, has four dense blocks with varying numbers of layers i.e., [6, 12, 24, 16]. The DenseNet-121 architecture was designed by Gao Huang (Huang et al., 2018). NasNet-Mobile is a CNN that explores an architectural building block on a small dataset and then assigns the block to a larger dataset, training the network by adding more layers into the block (Zhou & Diamos, 2018). Inception-V3 is a CNN model that mainly focuses on utilizing less computational power (Szegedy et al., 2016). It is a 48-layer deep network that stacks 11 inception modules. Each module consists of pooling layers and convolutional filters with rectified linear units as the activation function. Xception CNN stands for an extreme version of inception with a modified depth-separable convolution. The Xception architecture has 36 convolutional layers, with the initial two layers of convolution followed by depth-separable layers and fully connected layers (Chollet, 2017). MobileNet-V2 is a CNN that uses depth-wise separable convolution as efficient building blocks. It is a 53-layer deep network with 52 convolution layers and a fully connected layer (Sandler et al., 2018).

In this study, the bottom-up pathway was constructed using Fast RCNN-ResNet-50, with each of its five convolution modules (C1 through C5) containing multiple convolution layers (Yang et al. 2019). The spatial dimension was halved at each level from C1 to C5, reducing the C5 channel depth to 256-d (P5). Moreover, the initial feature map layer for object prediction was created using a  $1 \times 1$  convolution filter. The nearest neighbor upsampling technique was used in a top-down manner to upsample the preceding layer by a factor of two. The pixel-by-pixel  $1 \times 1$  filtered C4 and the upsampled P5 were combined to form P4. The same process was used to create P3 and P2. The network used its built-in multi-scale

pyramidal structure of deep convolutional networks to construct feature pyramids and generate autonomous predictions at different levels (P2, P3, P4, P5, and P6) for multiscale object detection. The general characteristics and structural details of all the CNN models are depicted in Tables 2 and 3. All models were trained on the Imagenet dataset (Neethi & Raviraj, 2024).

### 3.3. Performance Evaluation of Deep Learning Models

To compare the performance of the CNN models, 10 performance indices were calculated as follows:

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total Sample}} \quad (1)$$

$$\text{Loss} = \frac{-1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} * \log(p_{ij}) \quad (2)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (3)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (4)$$

$$F1 - \text{Score} = 2 * \frac{1}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} \quad (5)$$

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (6)$$

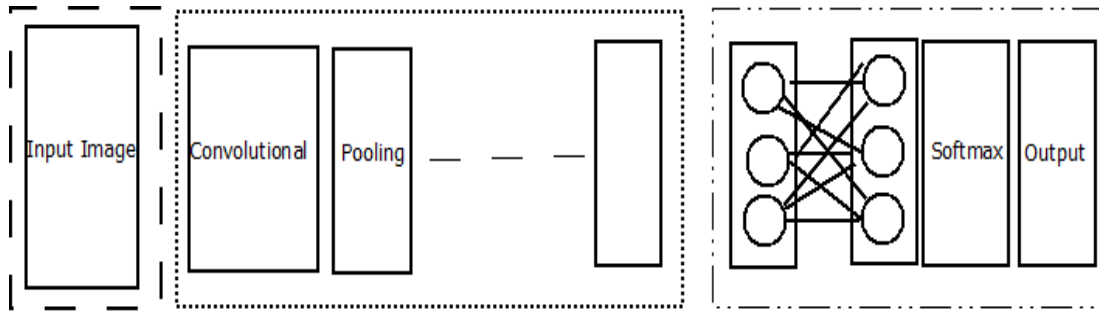
$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Positive} + \text{False Positive}} \quad (7)$$

$$\text{Positive Prediction Value (PPV)} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (8)$$

$$\text{Negative Prediction Value (PPV)} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Negative}} \quad (9)$$

$$\text{Cohen Kappa} = \frac{P_0 - P_e}{1 - P_e} \quad (10)$$





**Fig. 3.** Overview of the deep convolutional model.

**Table 2.** Characteristics of convolutional neural networks used.

Network	Depth	Parameters	Trained parameters	Input layer size	Output layer size	Epochs
AlexNet	8	21,619,464	32,776	224 by 224	7 by 7	100
VGG-16	16	14,718,792	4,104	224 by 224	7 by 7	100
VGG-19	19	20,028,488	4,104	224 by 224	7 by 7	100
FRCNN ResNet-50	50	23,604,104	16,392	224 by 224	7 by 7	100
ResNet-101	101	42,674,568	16,392	224 by 224	7 by 7	100
Inception-V3	159	21,819,176	16,392	224 by 224	7 by 7	100
NasNet-Mobile	-	4,278,172	8,456	224 by 224	7 by 7	100
MobileNet-V2	53	2,268,232	10,248	224 by 224	7 by 7	100
DensNet-121	121	7,045,704	8,200	224 by 224	7 by 7	100
Xception	71	20,877,872	16,392	224 by 224	7 by 7	100

**Table 3.** Structural details of the convolutional neural networks.

Network	Number of convolutional layers	Number of fully connected layers	Pooling	Softmax layer	Filters
AlexNet	5	3	3	1	256
VGG-16	13	3	5	1	512
VGG-19	16	3	5	1	512
FRCNN ResNet-50	16 (residual blocks)	1	1	1	2048
ResNet-101	33 (residual blocks)	1	1	1	2048
DensNet-121	4 (dense blocks)	1	1	1	1024
NasNet-Mobile	-	1	1	1	1056
Inception-V3	11 (inception blocks)	1	1	1	2048
Xception	36	1	1	1	2048
MobileNet-V2	7 (bottleneck)	1	1	1	1024

In this study, positive and negative classes were assigned to mango fruit and non-mango fruit regions. Hence, true positive and true negative represent the number of correctly diagnosed mango and non-mango regions, respectively. False positive and false negative represent the number of incorrectly diagnosed mango and non-mango regions, respectively. To evaluate the overall performance of the CNN model, the area under the receiver operating characteristic curve (Roc-AUC) and Cohen's kappa scores were calculated.

#### 4. Results

The performance of all 10 CNN model is depicted in Table 4. In this study, the training dataset consists of 4,000 images, and the validation dataset consists of 800 images. The networks were able to classify mango and non-mango regions of mango fruit crops, with a Roc-AUC score in the range of 0.96 to 0.99. The MobileNet-V2 CNN model outperformed all the other models. The performance of all the 10

CNN models was measured using accuracy, loss, Roc-AUC score, precision, recall, F1-score, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and Cohen's Kappa performance measure. The best model is considered based on the overall performance of all these measures.

The proposed FRCNN ResNet-50 network achieved a RoC-AUC score of 0.999 (accuracy: 0.9863; loss: 0.07; precision: 0.99; recall: 0.99; F1-score: 0.99; sensitivity: 0.94; specificity: 0.90; PPV: 0.96; NPV: 0.95; Cohen's Kappa: 0.98) in the training dataset and a RoC-AUC score of 0.989 (accuracy: 0.9815; loss: 0.04; precision: 0.98; recall: 0.98; F1-score: 0.98; sensitivity: 0.92; specificity: 0.87; PPV: 0.88; NPV: 0.97; Cohen's Kappa: 0.97) in the validation dataset. The time taken for each epoch was 2 seconds per step and 119 iterations per epoch.

The MobileNet-V2 network also achieved better results after the FRCNN ResNet-50, with a RoC-AUC score of 0.98 (accuracy: 0.97; loss: 0.09; precision:

**Table 4.** Overall performance of all the convolutional model.

Model	Data	Acc	Loss	Roc-AUC score	Pre	Rec	F1	Cohen's kappa	Sen	Spe*	PPV	NPV
AlexNet	Train	0.8245	0.4532	0.9657	0.85	0.82	0.81	0.8223	0.87	0.79	0.8901	0.6785
	Val	0.80	0.4328	0.9566	0.83	0.82	0.82	0.8103	0.9012	0.8345	0.9432	0.7012
NasNet-Mobile	Train	0.9356	0.1695	0.9980	0.96	0.95	0.95	0.9366	0.94	0.76	0.8293	0.9206
	Val	0.9275	0.1410	0.9977	0.95	0.94	0.94	0.9285	0.97	0.64	0.9552	0.7293
DenseNet	Train	0.9497	0.1588	0.9980	0.96	0.95	0.96	0.9393	0.9601	0.6765	0.8562	0.8943
	Val	0.945	0.1354	0.9983	0.96	0.95	0.94	0.9371	0.99	0.6765	0.9833	0.7071
VGG-16	Train	0.8986	0.2764	0.9934	0.90	0.89	0.90	0.87	0.8909	0.7209	0.7002	0.9010
	Val	0.8862	0.2918	0.9903	0.90	0.89	0.89	0.8783	0.91	0.67	0.8815	0.7338
VGG-19	Train	0.8844	0.3631	0.9905	0.88	0.87	0.87	0.934	0.9064	0.6395	0.8776	0.7076
	Val	0.8687	0.3542	0.991	0.88	0.87	0.87	0.901	0.97	0.55	0.6830	0.9482
Inception-V3	Train	0.9505	0.1170	0.9993	0.97	0.94	0.95	0.9518	0.9878	0.5851	0.9442	0.8716
	Val	0.946	0.1324	0.9991	0.96	0.95	0.95	0.9385	1.0	0.58	1.0	0.71
Xception	Train	0.7110	0.8199	0.9674	0.76	0.61	0.71	0.6444	0.98	0.103	0.8235	0.7586
	Val	0.5912	0.8982	0.9526	0.71	0.59	0.52	0.5328	0.96	0.06	0.8571	0.5052
ResNet-101	Train	0.96	0.1337	0.9987	0.97	0.96	0.97	0.9518	0.7901	0.9619	0.8791	0.9289
	Val	0.963	0.1096	0.9987	0.97	0.96	0.96	0.9585	0.7605	0.99	0.9870	0.8048
MobileNetV2	Train	0.9371	0.1942	0.9974	0.95	0.93	0.94	0.9238	0.9688	0.6024	0.8714	0.8742
	Val	0.9312	0.1741	0.9979	0.95	0.93	0.93	0.9214	0.99	0.6024	0.9821	0.6875
FRCNN-ResNet 50	Train	0.9715	0.0912	0.9993	0.98	0.97	0.98	0.9657	0.9844	0.7802	0.9461	0.9274
	Val	0.9685	0.0797	0.9996	0.97	0.97	0.97	0.9642	0.99	0.76	0.9870	0.8048

Abbreviations: Acc: Accuracy; F1: F1-Score; NPV: Negative predictive value; PPV: Positive predictive value; Pre: Precision; Rec: Recall; Sen: Sensitivity; Spe: Specificity; Val: Validation.

0.98; recall: 0.98; F1-score: 0.98; sensitivity: 0.92; specificity: 0.95; PPV: 0.96; NPV: 0.95; Cohen's Kappa: 0.98) and a RoC-AUC score of 0.989 (accuracy: 0.98; loss: 0.05; precision: 0.98; recall: 0.98; F1-score: 0.98; sensitivity: 0.92; specificity: 0.96; PPV: 0.95; NPV: 0.89; Cohen's Kappa: 0.97) in the training and validation datasets, respectively. The time taken for each epoch was 6 seconds per step, with 119 iterations per epoch.

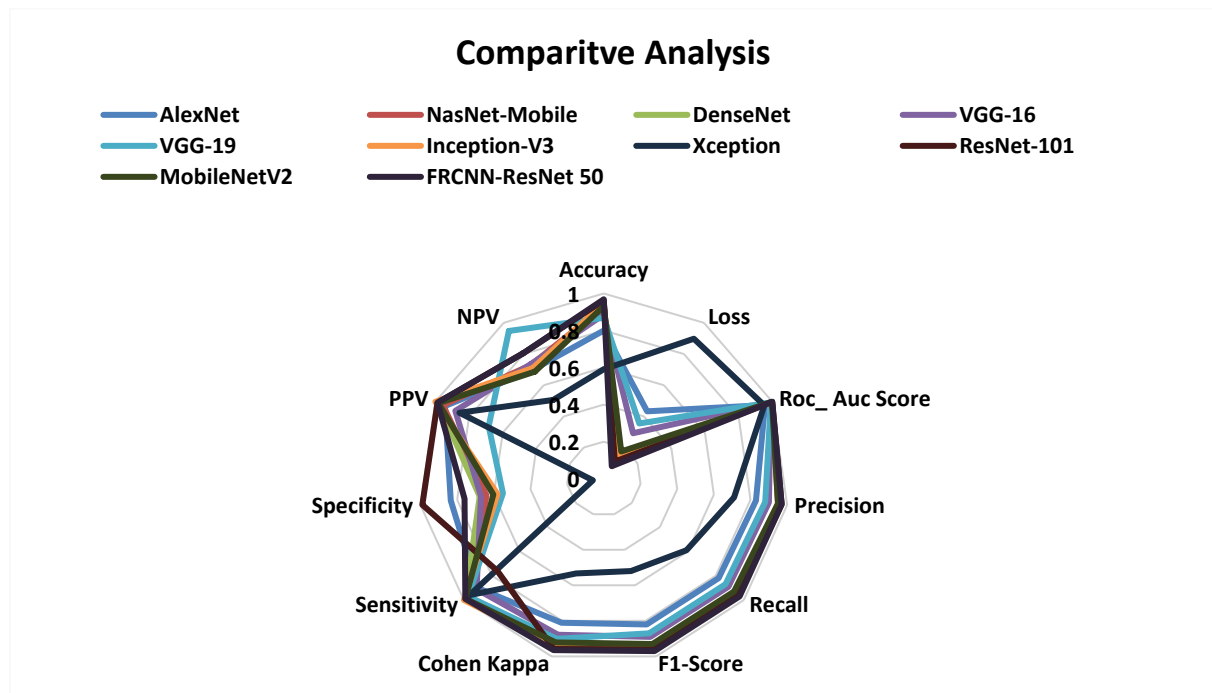
The ResNet-101 network achieved a RoC-AUC score of 0.96 and an accuracy of 0.71, with overall performance much lower compared to other networks. The time taken for each epoch was 11 seconds per step, with 119 iterations per epoch. The NasNet-Mobile network also achieved good results, close to the Xception network. DenseNet, Vgg-19, Inception-V3, and AlexNet networks performed well in classifying healthy and diseased leaves, with accuracy in the range of 0.91–0.94. VGG-16 achieved an accuracy of 0.88.

Fig. 4 shows the overall performance of all the networks using a radar plot. The training and validation process is shown in Fig. 5. A confusion matrix was calculated to simplify the understanding of the performance of all the networks on both training and validation datasets, as shown in Fig. 6. In the confusion matrix, the eight classes of mango crops were represented by class 0 to class 7: class 0 — Mallika, class 1—Raspuri, class 2—Arka Arun, class 3—Malgoa, class 4—Badam, class 5—Kesar, class

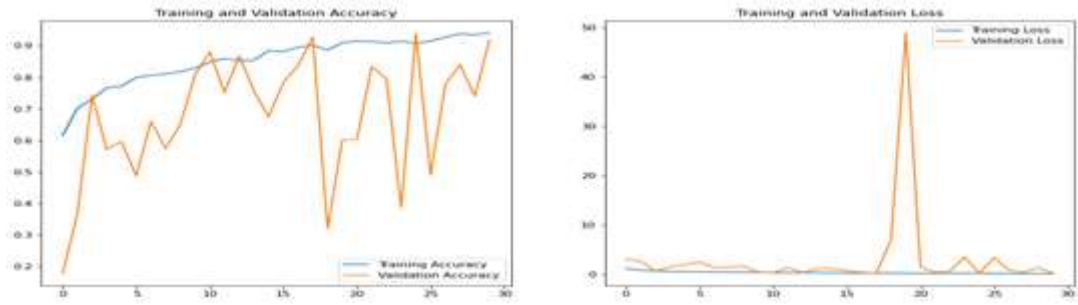
6—Sindoora, and class 7—Alphanso (Table 1). By observing all the model's confusion matrices, it can be noted that the major misclassifications were between class 0 and class 1, i.e., Mallika and Raspuri. The model failed to classify the images captured under field conditions due to the presence of other parts of the plant, the ground, and other background objects. Even after tuning hyperparameters, the results did not show considerable variance. The reputable results were achieved for 100 epochs. The classification accuracy of each mango fruit crop class for all the convolutional models is depicted in Table 5, and the respective chart is shown in Fig. 7. Receiver operating characteristic curve for each class are plotted and presented in Fig. 8 for all the convolutional models.

## 5. Discussion

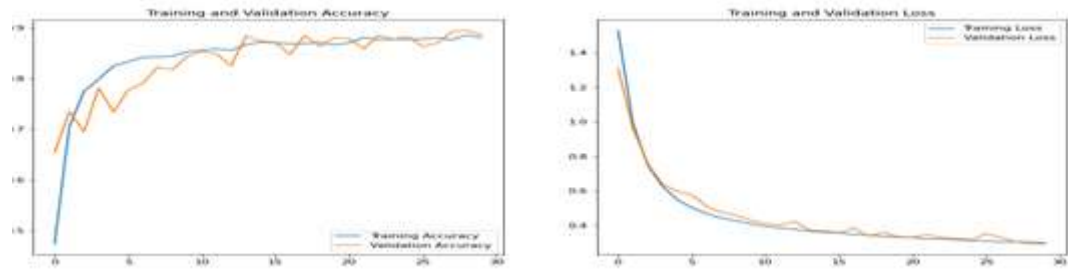
In this study, an FRCNN model was proposed and compared with well known CNNs to estimate mango fruit crop yields. The results showed that proposed method could efficiently estimate mango crop yield with the highest accuracy. Although the proposed model and MobileNet-V2 both gave the same accuracy results, the proposed model could diagnose plant disease with higher sensitivity, lower specificity,



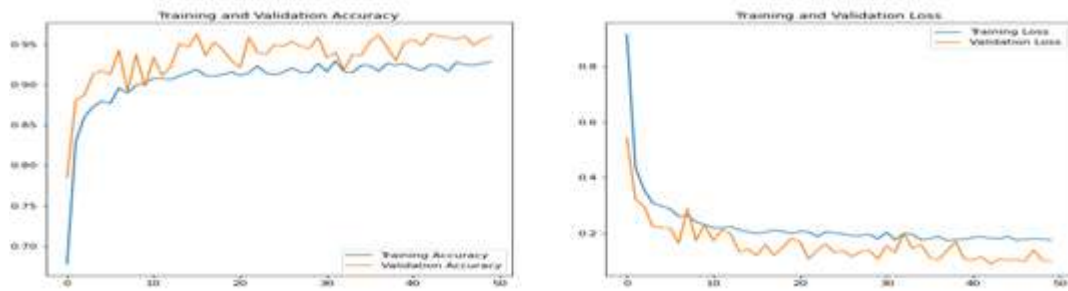
**Fig. 4.** Radar plot of the 10 individual networks. Abbreviations: NPV: Negative predictive value; PPV: Positive predictive value; Roc\_Auc: Area under the receiver operating characteristic curve.



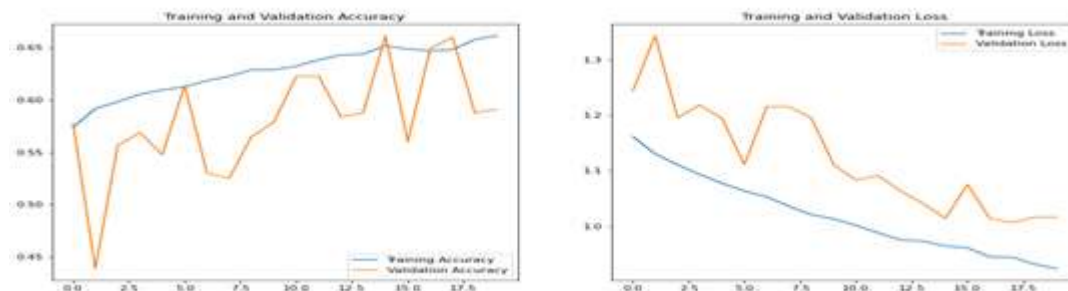
1. AlexNet



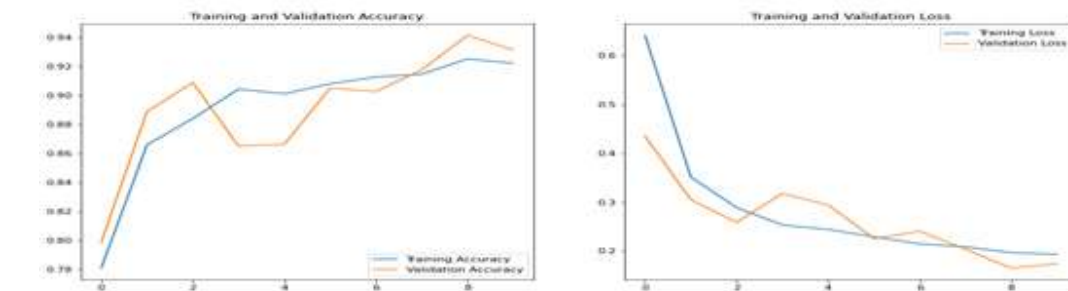
2. Vgg16



3. Vgg19

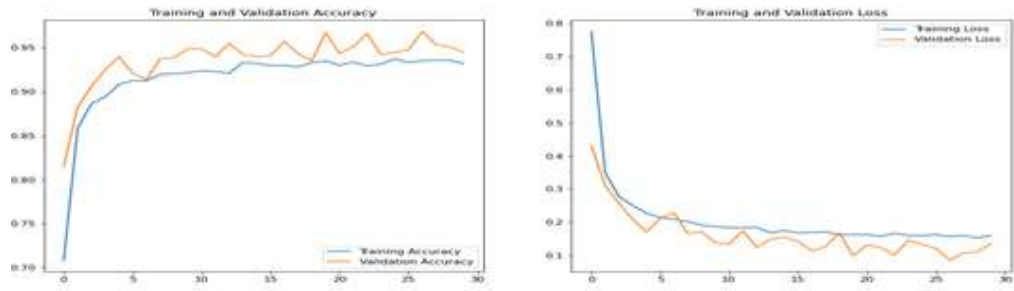


4. ResNet-101

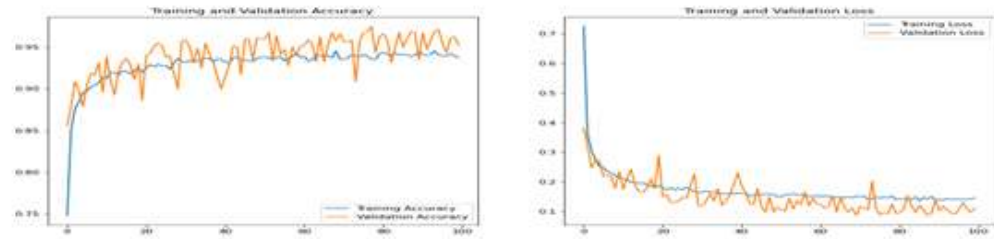


5. DesnseNet

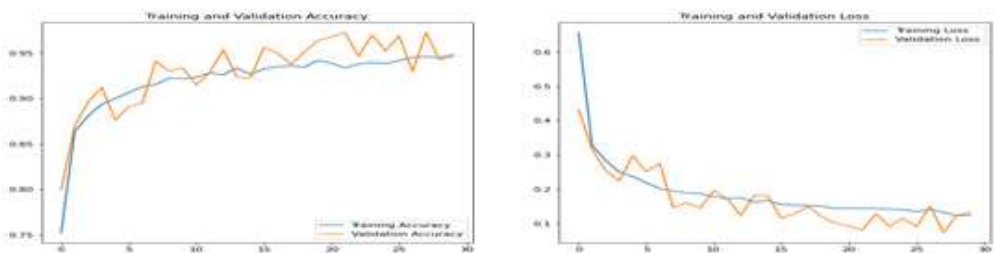
(Fig. 5)



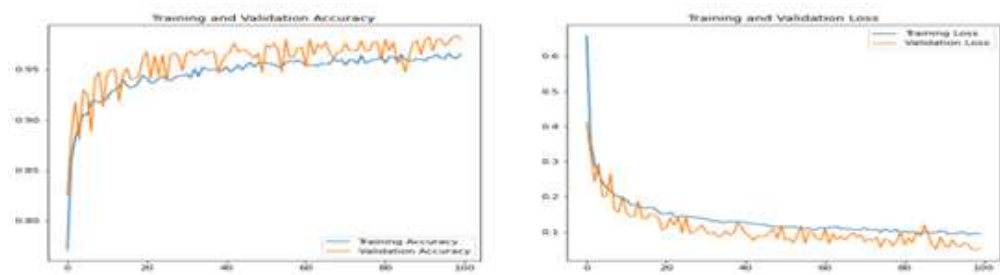
6. NasNet-Mobile



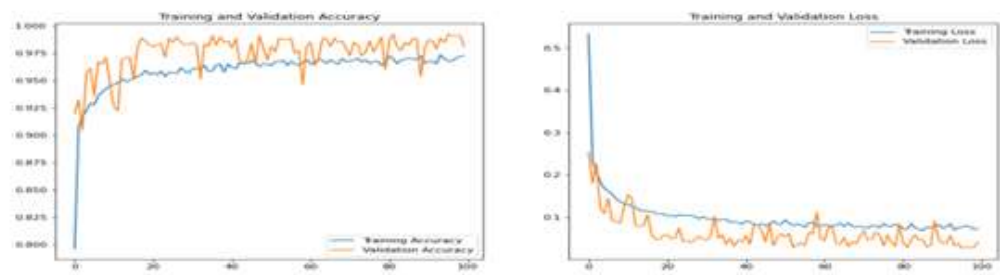
7. Inception-V3



8. Xception

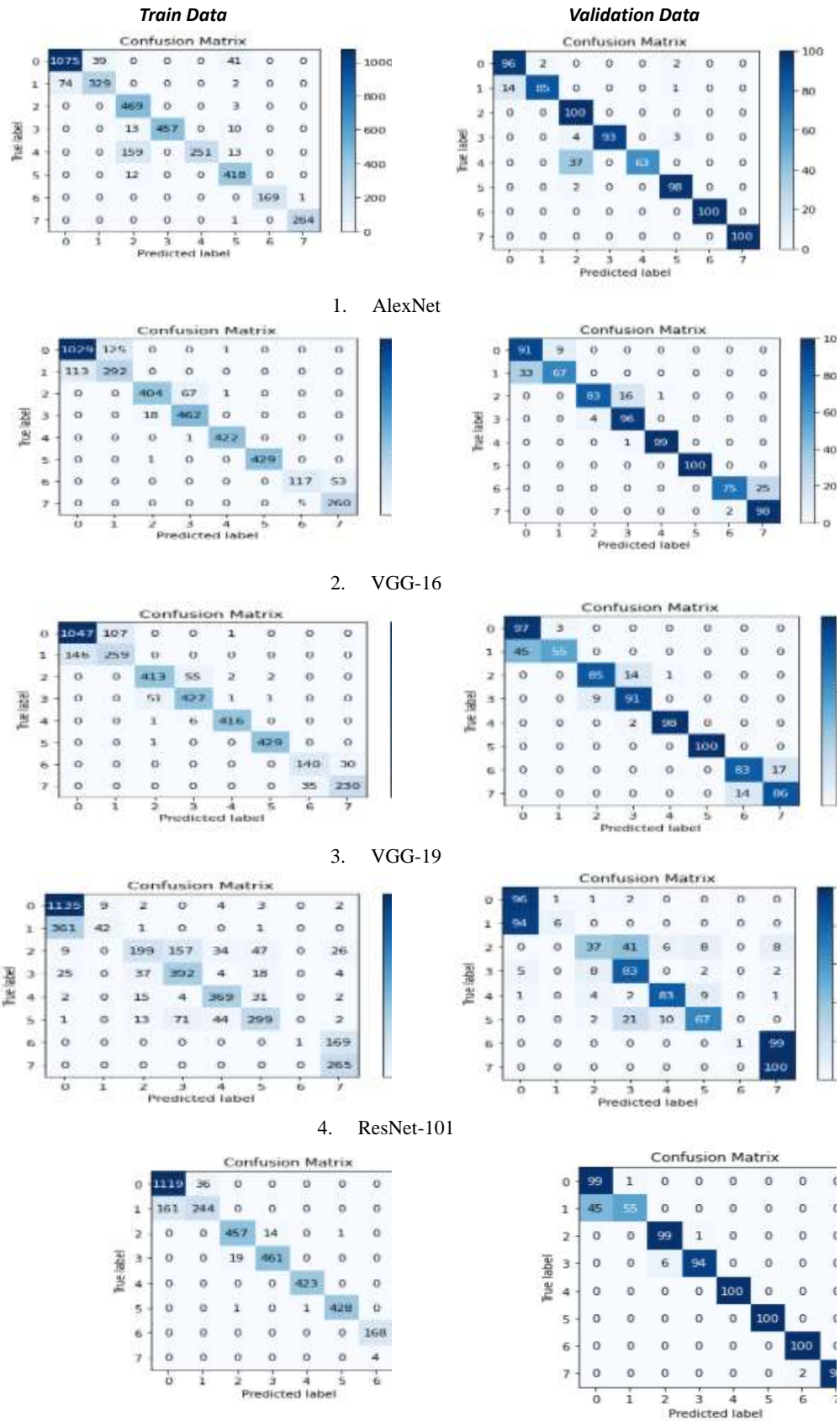


9. MobileNet-V2



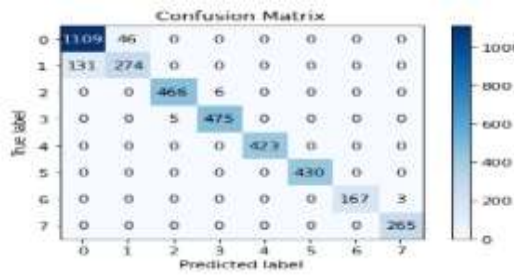
10. FRCNN ResNet-50

Fig. 5. Accuracy (right) and loss (left) plots of 10 convolutional neural networks for training and validation datasets.

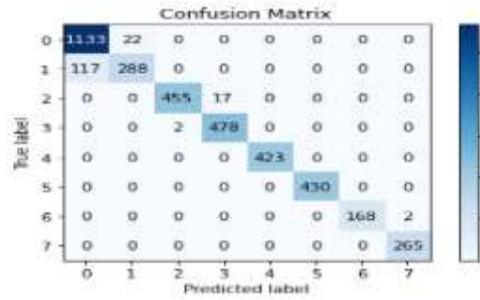
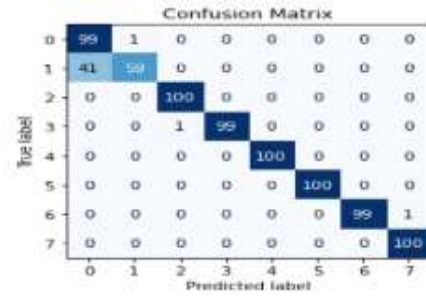


(Fig. 6)

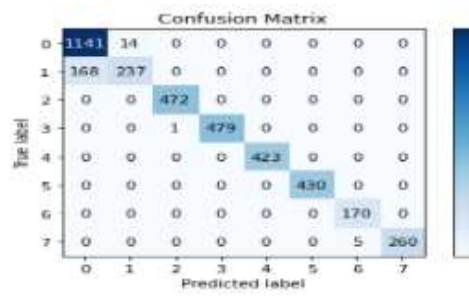
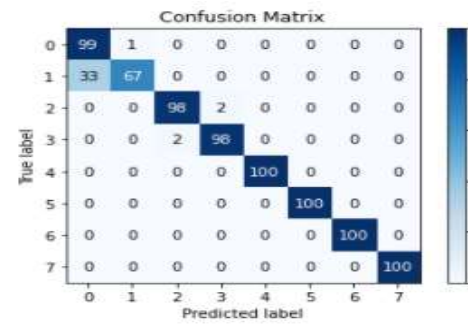




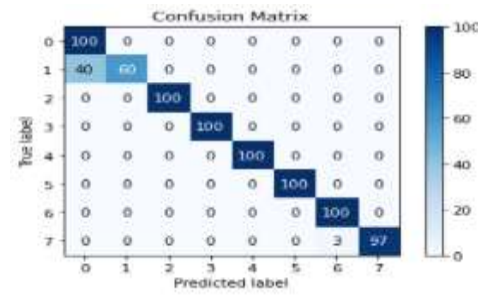
6. NasNet-Mobile



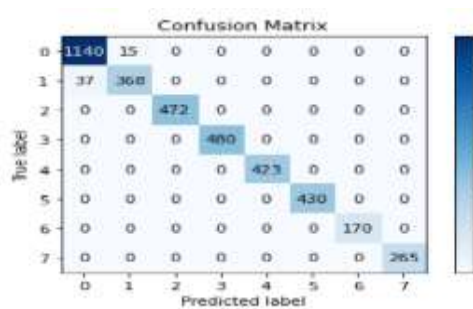
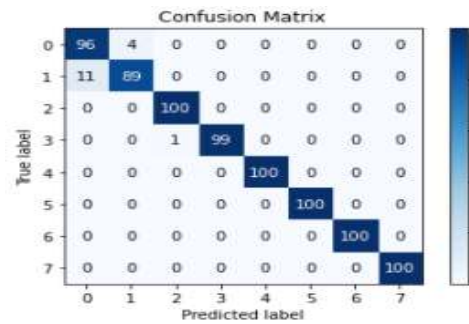
7. Inception-V3



8. Xception



9. MobileNet-V2



10. FRCNN-ResNet-50

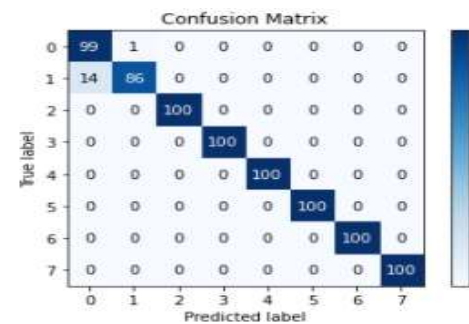


Fig. 6. Confusion matrices of 10 convolutional neural networks for training and validation datasets.



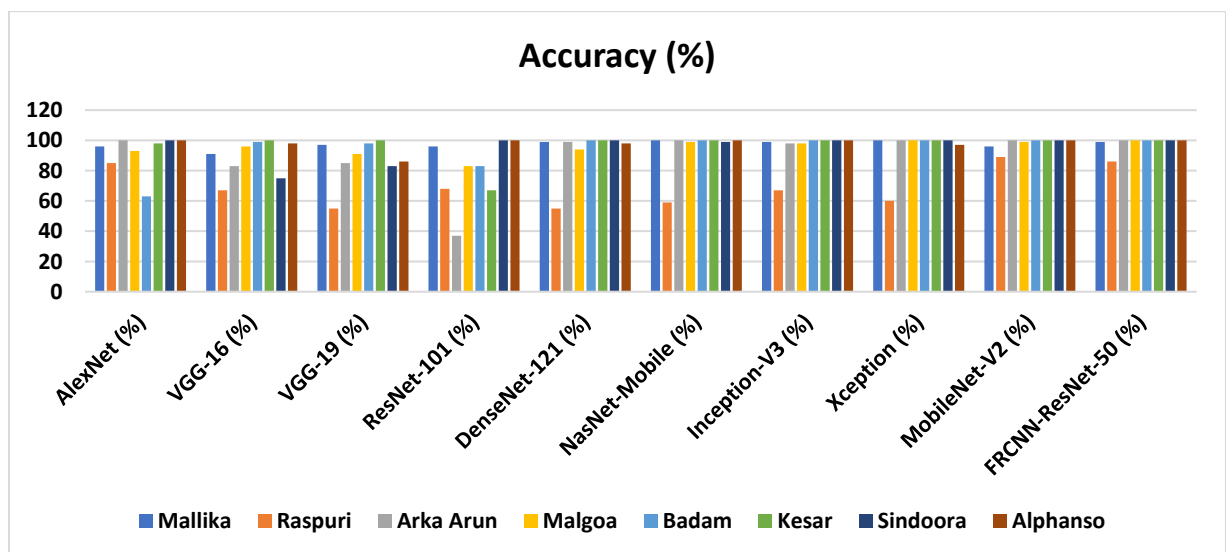
**Table 5.** Classification accuracy of the fruit crop leaves dataset for each class.

Mango Fruit Culti- vars	Accuracy (%)	AlexNet (%)	VGG-16 (%)	VGG-19 (%)	ResNet-101 (%)	DenseNet-121 (%)	NasNet-Mobile (%)	Inception-V3 (%)	Xception (%)	MobileNet-V2 (%)	FRCNN-ResNet-50 (%)
Mallika	100	96	91	97	96	99	100	99	100	96	99
Raspuri	100	85	67	55	68	55	59	67	60	89	86
Arka Arun	100	100	83	85	37	99	100	98	100	100	100
Malgoa	100	93	96	91	83	94	99	98	100	99	100
Badam	100	63	99	98	83	100	100	100	100	100	100
Kesar	100	98	100	100	67	100	100	100	100	100	100
Sindoora	100	100	75	83	100	100	99	100	100	100	100
Alphanso	100	100	98	86	100	98	100	100	97	100	100
Overall Accuracy of the Model	100	91.3	88.6	86.87	59.1	93.1	94.5	95.25	94.6	98.0	98.85

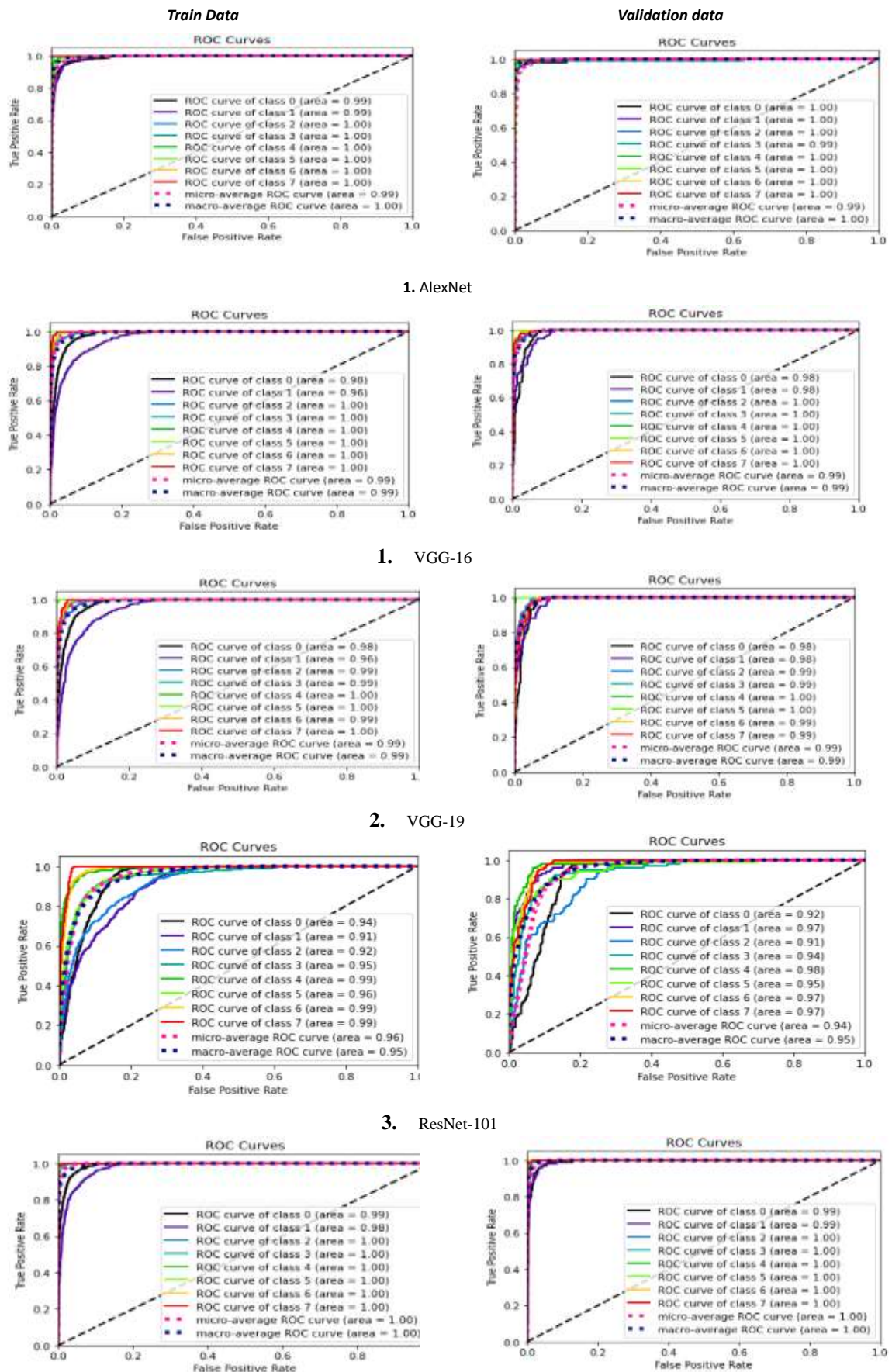
and greater PPV when compared to the MobileNet-V2 network. The proposed model has a depth convolution structure, and the network becomes lighter as it goes deeper. Its inverted residual block helps improve efficiency and boosts the robustness of the model.

The MobileNet-V2 network achieved an accuracy of 85.12% using 4,000 images of five different classes of fruits. In one study, fruit images were used to detect the diseases present in the images,

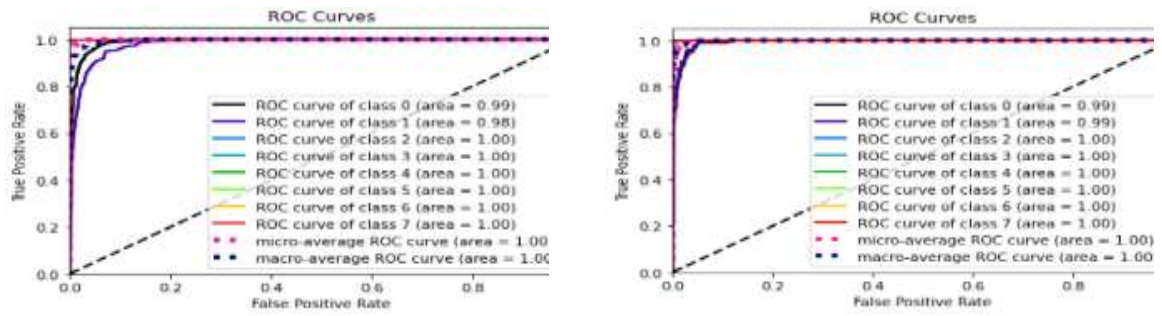
and the results were compared with MobileNetV1, InceptionV3, and DenseNet121 (Xiang et al., 2019). The classification accuracy of all eight classes of fruit crops for MobileNet-V2 is as follows: class 0—diseased coffee, 99%; class 1—grape black rot, 86%; class 2—grape esca, 100%; class 3—grape healthy, 100%; class 4—grape leaf blight, 100%; class 5—healthy coffee, 100%; class 6 —diseased mango, 100%; class 7—healthy mango, 100%. A



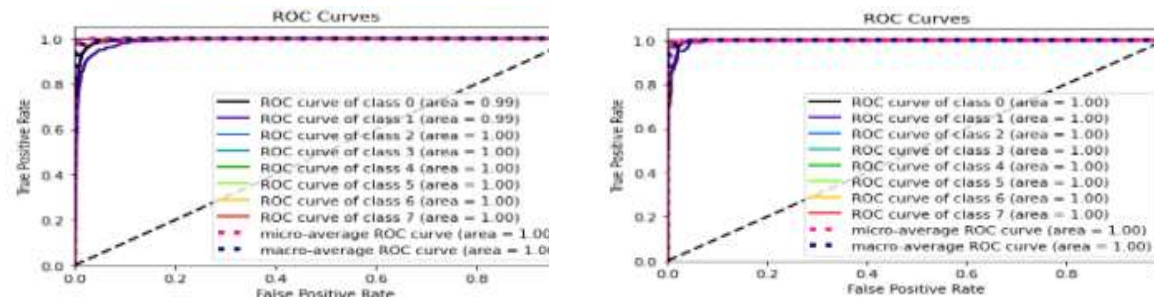
**Fig. 7.** Classification accuracy chart of the fruit crop leaves dataset for each class.



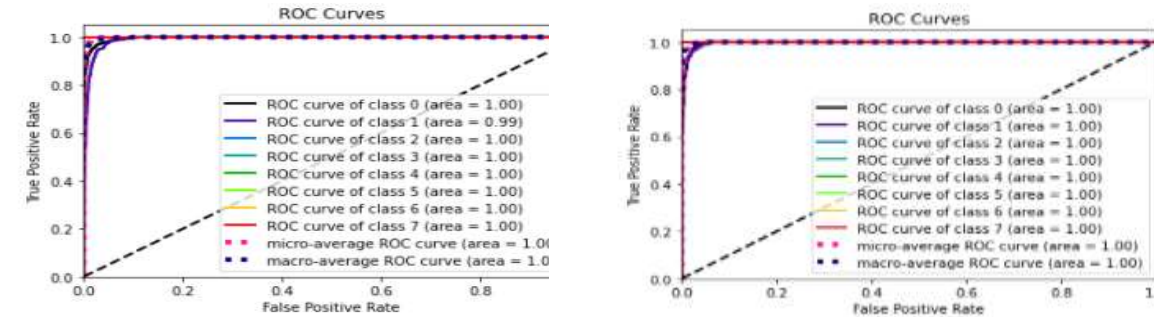
(Fig. 8)



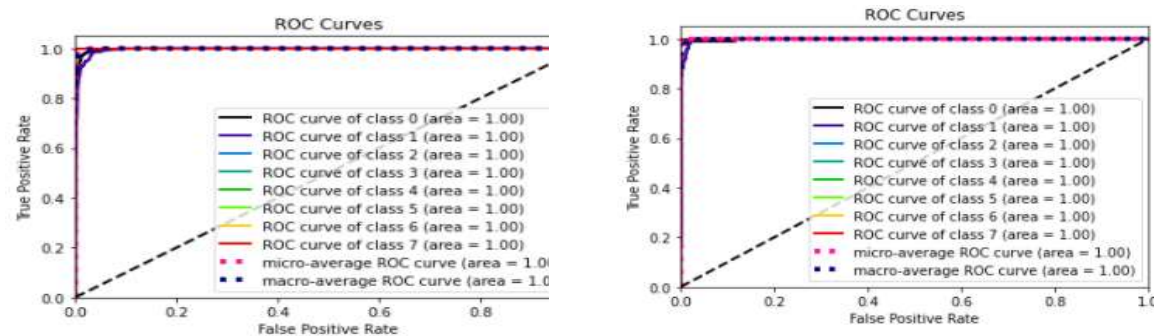
## 5. NasNet-Mobile



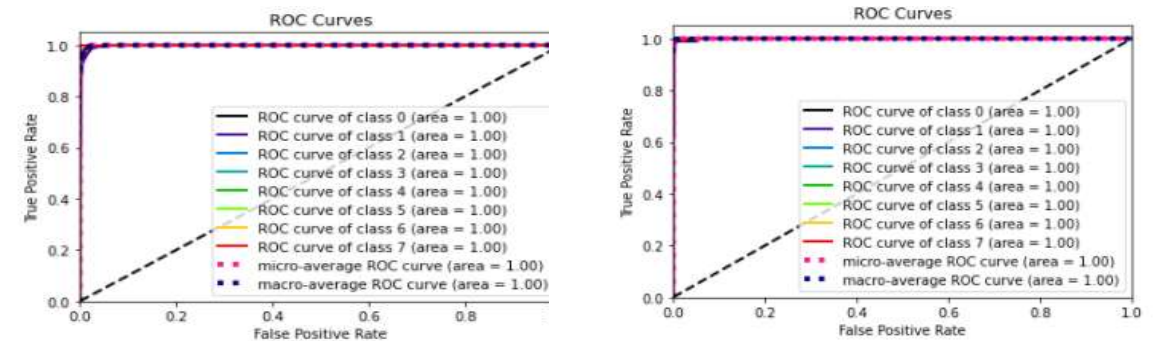
## 6. Inception-V3



## 7. Xception



## 8. MobileNet-V2



## 9. FRCNN ResNet-50

Fig. 8. Receiver operating characteristic curve of 10 convolutional neural networks for training and validation datasets.

convolutional network was constructed to classify 22 different plant disease classes. The classification accuracy of the network ranged from 33% to 98%, with an average accuracy of 86.2%. The network failed to classify a few classes due to the smaller number of training samples (Dyrmann et al., 2016). The Tomato Diseases and Pests dataset, which contains challenging images of diseases and pests, was used. This dataset includes several inter-and extra-class variations, such as infection status and location in the plant. VGGNet and ResNet were combined to form deep learning meta-architectures formed to train the tomato images, achieving an accuracy of 80% (Fuentes et al., 2017; Saleem et al., 2019). Hybrid convolutional models have been shown to help to detect plant diseases, offering a new direction for plant disease detection (Punam & Gole, 2021), which could be adopted in future studies. Better data retrieval systems assist in retrieving suitable plant disease datasets, and texture techniques with machine learning algorithms can help select the best dataset (Dhingra & Bansal, 2020). The limitation of this study is that the networks could not classify images captured under field conditions, and pre-processing was required for images from field conditions. Additionally, the dataset size should be increased, and more images need to be used for training.

The study provides the following contributions for the literature:

- Introduces a novel FRCNN model that enhances the estimation of mango crop yield and the diagnosis of plant diseases, contributing to the body of knowledge on agricultural applications of CNNs.
- Contributes to the literature by demonstrating how an FRCNN model can achieve higher sensitivity and PPV in plant disease diagnosis compared to other models.
- The study contributes by identifying key limitations in current approaches, such as the need for larger datasets and better handling of field-captured images, which can guide future developments in the field.
- The study suggests that hybrid convolutional models and improved data retrieval systems could significantly advance the field of plant disease detection.

## 6. Conclusion

An efficient approach to estimating the harvest of mango fruits was proposed in this work. The dataset used to train the proposed approach consisted of

images of eight different mango varieties. Nine pre-trained models, of which only the final layer was altered, were compared with the proposed FRCNN model. Each model's classification performance was improved by fine-tuning the model. To improve accuracy and reduce error rates, the models were subjected to validation and optimization. The proposed model's accuracy was determined to be 98.85%. The results demonstrated that pre-trained models were unable to accurately estimate mango yield. As a result, a novel CNN architecture was proposed and applied in this study.

Adapting the CNN model presented in this work improves mango fruit yield estimation performance. This method will assist individuals with limited knowledge of mango crop yield estimation, as it can be difficult to receive proper guidance from agriculturists for manual variety determination and categorization. At the same time, it will enable precise, quick, and reliable classification. By utilizing various deep learning techniques, we hope to expand the scope of this study and include more types and data in the future.

## Acknowledgments

The support by GSSS Institute of Engineering and Technology for Women, Affiliated to Visvesvaraya Technological University (VTU), Belagavi, India.

## References

- Ali, H., Lali, M.I., Nawaz, M.Z., Sharif, M., & Saleem, B.A. (2017). Symptom-based automated detection of citrus diseases using color histogram and textural descriptors. *Computers and Electronics in Agriculture*, 138, 92–104.
- Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A., & Stefanović, D. (2019). Solving current limitations of deep learning-based approaches for plant disease detection. *Symmetry*, 11, 939.
- Atila, Ü., Uçar, M., Akyol, K., & Uçar, E. (2021). Plant leaf disease classification using EfficientNet deep learning model. *Ecological Informatics*, 61, 101182.
- Bedi, P., & Gole, P. (2021). Plant disease detection using a hybrid model based on convolutional autoencoder and convolutional neural network. *Artificial Intelligence in Agriculture*, 5, 90–101.
- Brahimi, M., Mahmoudi, S., Boukhalfa, K., &



- Moussaoui, A. (2019). Deep interpretable architecture for plant diseases classification. 2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), 111–116.
- Chollet, F. (2016). Xception: Deep learning with depthwise separable convolutions. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1800–1807.
- Dhingra, S., & Bansal, P. (2020). Employing divergent machine learning classifiers to upgrade the preciseness of image retrieval systems. *Cybernetics and Information Technologies*, 20, 75–85.
- Fine, T.L., Lauritzen, S.L., Jordan, M., Lawless, J., & Nair, V. (1999). Feedforward neural network methodology. Springer-Verlag, Berlin, Heidelberg.
- Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17(9), 2022.
- Grinblat, G.L., Uzal, L.C., Larese, M.G., & Granitto, P.M. (2016). Deep learning for plant identification using vein morphological patterns. *Computers and Electronics in Agriculture*, 127, 418–424.
- Gyires-Tóth, B.P., Osváth, M., Papp, D., & Szűcs, G. (2019). Deep learning for plant classification and content-based image retrieval. *Cybernetics and Information Technologies*, 19(1), 88–100.
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778.
- Heredia, I., Iglesias, L.L., Vykozlov, V., & Orviz, P. (2019). Plants classification engine. DIGITAL.CSIC.
- Huang, G., Liu, Z., & Weinberger, K.Q. (2016). Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2261–2269.
- Khan, A., Sohail, A., Zahoor, U., & Qureshi, A.S. (2019). A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, 53, 5455–5516.
- Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60, 84–90.
- Mohanty, S.P., Hughes, D.P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Computers and Electronics in Agriculture*, 7, 1419.
- Neethi, M.V., & Raviraj, P. (2024). Fast region-based convolutional neural network ResNet-50 model for on-tree mango fruit yield estimation. *Indonesian Journal of Electrical Engineering and Computer Science*, 33, 1084.
- Sandler, M., Howard, A.G., Zhu, M., Zhmoginov, A., & Chen, L. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 4510–4520.
- Saleem, M.H., Potgieter, J., & Arif, K.M. (2019). Plant disease detection and classification by deep learning. *Plants*, 8.
- Shi, Y., Huang, W., Luo, J., Huang, L., & Zhou, X. (2017). Detection and discrimination of pests and diseases in winter wheat based on spectral indices and kernel discriminant analysis. *Computers and Electronics in Agriculture*, 141, 171–180.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556.
- Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks-based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience*, 6, 1–11.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2015). Rethinking the inception architecture for computer vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2818–2826.
- Xiang, Q., Wang, X., Li, R., Zhang, G., Lai, J., & Hu, Q. (2019). Fruit image classification based on MobileNetV2 with transfer learning technique. *Proceedings of the 3rd International Conference on Computer Science and Application Engineering*.
- Yang, K., Qinami, K., Fei-Fei, L., Deng, J., & Russakovsky, O. (2020). Towards fairer datasets: Filtering and balancing the distribution of the people subtree in the ImageNet hierarchy. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT '20)\**, 547–558.
- Yang, J., Sun, J., Ren, Y., Li, S., Ding, S., & Hu, J. (2023). GACP: Graph neural networks with ARMA filters and a parallel CNN for hyperspectral image classification. *International*

- Journal of Digital Earth, 16(1), 1770–1800.
- Yang, X., & Tingwei, G. (2017). Machine learning in plant disease research. *European Journal of Biomedical Research*, 6–9.
- Zhang, S., Wu, X., You, Z., & Zhang, L. (2017). Leaf image-based cucumber disease recognition using sparse representation classification. *Computers and Electronics in Agriculture*, 134, 135–141.
- Zhang, X., Han, L., Dong, Y., Shi, Y., Huang, W., Han, L., González-Moreno, P., Ma, H., Ye, H., & Sobeih, T. (2019). A Zhou, D.X. (2020). Theory of deep convolutional neural networks: Downsampling. *Neural Networks*, 124, 319–327.
- deep learning-based approach for automated yellow rust disease detection from high-resolution hyperspectral UAV images. *Remote Sensing*, 11(13), 1554.
- Zhou, Y., & Damos, G. (2018). Neural architect: A multi-objective neural architecture search with performance prediction. *arXiv preprint arXiv:1804.09081*.

He is currently guiding Ph.D. research scholars in the areas of image processing, pervasive and cloud computing, bio-inspired algorithms, and robotics. He has received a project grant of Rs.5 Lakhs from the VGST, Govt. of Karnataka, for the “Underwater Robotic Fish for Surveillance and Pollution monitoring.” He has received the awards and recognitions such as the “Rhasriya Gaurav Award, 2015,” “Shri P.K. das Memorial Best Faculty Award, 2012,” and “Young Achiever Award, 2016.” He can be contacted at the email address: [raviraj@gsss.edu.in](mailto:raviraj@gsss.edu.in).

## AUTHOR BIOGRAPHIES



**Neethi M. V.** completed her doctorate degree in Computer Science and Engineering in the area of machine learning and received an M.Tech. degree in Computer Cognition

Technology from University of Mysore, India, in 2015. She also received her B.E. from Visvesvaraya Technological University, India, in 2013. She is currently an assistant professor in the Department of Computer Science and Engineering—Data Science at ATMECE, VTU, Karnataka, India. Her research areas include machine learning, data mining, and image processing. She can be contacted at the following email addresses: [neethi.m@gmail.com](mailto:neethi.m@gmail.com) or [neethimv\\_cd@atme.edu.in](mailto:neethimv_cd@atme.edu.in).



**Raviraj P.** completed his doctorate degree in Computer Science and Engineering in the area of Image Processing. He holds the position of Director-IQAC and Professor in the

Department of Computer Science and Engineering at GSSS Institute of Engineering and Technology for Women, Mysore, Karnataka. He has 19 years of teaching and research experience. He has published more than 92 papers in international journals and conferences. Five research scholars have completed their Ph.D. under his guidance at various universities.

# Technology innovation of dryer machine based on sustainability automation systems to increase agel fiber production in handicraft SME

Khakam Maruf<sup>1,2</sup>, Rizal Justian Setiawan<sup>2,3,8\*</sup>, Darmono<sup>4</sup>, Syukri Fathudin Achmad Widodo<sup>2</sup>, Sumantri Sri Nugroho<sup>5</sup>, Nur Evirda Khosyati<sup>6</sup>, Nur Azizah<sup>7</sup>

<sup>1</sup>Industrial Engineering, Faculty of Engineering, Gadjah Mada University, Indonesia

<sup>2</sup>Mechanical Engineering Education, Faculty of Engineering, Yogyakarta State University, Republic of Indonesia

<sup>3</sup>Industrial Engineering and Management, College of Engineering, Yuan Ze University, Taiwan (R.O.C)

<sup>4</sup>Civil Engineering and Planning Education, Faculty of Engineering, Yogyakarta State University, Indonesia

<sup>5</sup>Balai Pengembangan Teknologi Tepat Guna (BPTTG), Disperindag DI, Yogyakarta, Indonesia

<sup>6</sup>Culinary Technology Education, Faculty of Engineering, Yogyakarta State University, Indonesia

<sup>7</sup>International Public Health, School of Public Health, China Medical University, Taiwan (R.O.C)

<sup>8</sup>Asia and China Studies, School of Law and Politics, National Chung Hsing University, Taiwan (R.O.C)

\*Corresponding author E-mail: rizaljustiansetiawan99@gmail.com

(Received 17 April 2024; Final version received 25 December 2024; Accepted 24 January 2025)

## Abstract

In Indonesia, natural fibers are extensively utilized as essential raw materials for various human needs. These natural fibers find significant application in handicrafts, particularly in small and medium enterprises (SME) within the handicraft industry. Agel, a prominent natural fiber, is obtained from drying gebang leaves and plays a vital role in the local economy of the Kulon Progo region in Yogyakarta, Indonesia. Currently, the production process of raw materials to become fiber relies on conventional methods, primarily sun-drying, which often gives rise to numerous challenges such as temperature fluctuations and weather dependencies. Moreover, the physical posture adopted by workers during the drying process is ergonomically unfavorable, as they must bend over repeatedly to turn the agel fibers being dried under the sun. After recognizing these issues, it becomes imperative to develop a sustainable dryer machine equipped with advanced technology that enhances productivity while prioritizing employee ergonomics. This study employs the research and development (R&D) method, which encompasses analysis, design, development, implementation, and evaluation stages. The primary objective of this research is to design, fabricate, and test a dryer machine utilizing a sustainability automation system integrated with Internet of Things (IoT). The outcome of this research is a dryer machine that can effectively dry agel leaves within a significantly reduced timeframe of 2–4 hours, with a maximum capacity of 10 kg per cycle. This achievement surpasses the conventional method, which typically takes 5–6 days to produce 10 kg of dried agel fibers.

**Keywords:** Agel Fiber, Crafts, Dryer Machine, Gebang Leaves; SME

## 1. Introduction

Indonesia is a country that is rich in natural biological resources to be utilized (Edy, 2023), Indonesia's natural resources and its ecosystem have an important

position and role to be utilized for life and national development (Albert et al., 2022; Nikijuluw, 2017). One of Indonesia's biological riches is plants that produce natural fiber. Natural fibers have many benefits and are widely used as raw materials needed by humans (Islam



& Mohammad, 2016; Marian et al., 2022). The need for natural fiber raw materials in Indonesia is quite high, this is proven by data on the value of imports of raw materials. Natural fibers are usually used as crafts, alternative additional materials, or as composite reinforcement (Kumar & Sekaran, 2014). Using natural fiber plants is considered more economical, easy to apply, corrosion resistant, and easy to obtain (Ali et al., 2021; Mochane et al., 2019; Zwawi, 2021). Along with the development of science and technology, the demand for natural fiber raw materials is increasing (Nurhaliza et al., 2023).

Agel fiber is a type of natural fiber that comes from the gebang leaves with the Latin name *Coryphatan* (Nugroho et al., 2018). The agel fiber is a promising natural fiber from the gebang plant with several advantages (Nuryanta et al., 2023). In Indonesia, this type of fiber is widely processed into various types of creative crafts, such as bags, decorations, wallets, and hats. In the process, agel fiber comes from raw gebang leaves which must go through 5 stages before it is ready to be woven. The initial stage is separating the stick from gebang leaves, then brushing to reduce the size, drying, dyeing, and finally being woven into raw material for crafts. The stages in the production of agel fiber raw materials are still using conventional methods.

The handicraft industry from agel fiber is one of the economic sectors driving local communities in the Sentolo District area, Kulon Progo, Yogyakarta (Widodo et al., 2020). Handicrafts made from agel fiber have an export target and are one of the products with quite high market interest. One of the agel fiber craft processing enterprises is located in Giling Hamlet, Tuksono, Sentolo, Kulon Progo, Yogyakarta. If the weather conditions are favorable, every month the enterprise is able to process gebang leaves and produce 150 kg of agel with an average of 5 kg of agel every day, and with 150 kg of agel, it is able to produce 300 to 500 handicrafts for one month. The average income earned for 1 month is IDR 9,000,000 or USD 591 (Exchange Rate USD 1 = IDR 15,227.24 on 10 August 2023). The raw material production process is still conducted conventionally by relying on sunlight, which often causes various problems, such as weather conditions, ergonomics issues, and longtime processes.

Working position and attitude during the drying process also has an unhealthy impact on body posture because when drying the workers must bend over to turn the agel that is being dried in the sun. The posture at work will indirectly change the shape and position

of the spine or cause injury and musculoskeletal disorders (MSDs) (Arnau et al., 2023; Bonfiglioli et al., 2022; Caponecchia et al., 2020). MSDs are disorders of the musculoskeletal system that are experienced by a person ranging from mild complaints to very painful (Irwan et al., 2023). If the muscles receive a static load in the long term, it will cause complaints in the form of damage to joints, ligaments, muscles, and other skeletal systems. Musculoskeletal disorders are one of the most common health problems experienced by workers (Kim, 2022; You et al., 2020).

Based on these problems, the researchers created Technology Innovation of Dryer Machine based on Sustainability Automation Systems to Increase Agel Fiber Production in Handicraft SMEs. The drying machine shortens the drying time of agel, which is 2–4 hours. The machine can also be used to dry agel fiber after the coloring process. This step makes it easier to produce agel fiber into handicrafts. The application of this machine makes the work position healthier because SME workers no longer must bend over to dry and turn the gel that is dried in the sun, this will be a solution in reducing the risk of MSDs.

The existence of an agel drying machine will help SME in increasing production without being constrained by the drying process, and their turnover is increasing compared to the conventional drying process with solar heat. The dryer machine has a monitoring system for temperature, humidity and drying time that has been integrated with an application on a smartphone so that the drying process of agel fiber can be monitored regularly. The application of the drying machine is expected to optimize the craft industry sector which can make SME businesses develop into sustainable businesses.

## 2. Literature Review

### 2.1. Relevant Previous Research

In this research, the fuel source uses liquefied petroleum gas (LPG). This was chosen based on conditions where LPG is easy to obtain and affordable to use in Indonesia. This system uses LPG as a heat source that is more controllable and uniform compared to drying using charcoal or solar power. Test results showed that this system has an energy efficiency of up to 90% and an exergy efficiency of up to 10%. This shows that this system effectively utilizes LPG gas input energy for the drying process (Benjamin et al., 2022). Apart from that, the drying system using LPG is more

efficient than conventional drying systems (Saparudin et al., 2021).

The second research obtained a recirculated batch-type corn drying machine with a capacity of 2 tons of corn per 8 hours. This machine uses a direct heating system with LPG fuel. Machine trials showed good performance with a drying rate of 1.55%/hour, heating efficiency of 92.03%, and drying efficiency of 88.96%. Overall, the test results showed that the design of the drying machine using LPG gas is quite optimal for drying corn on a production scale (Karyadi, 2019).

The research conducted is supported by several previous studies, where several studies that have been conducted state that electricity-based drying systems can increase the production of a business. In research conducted by Lasano (2018) stated that the use of microwaves for drying is an alternative drying method. This drying method is classified as efficient in terms of energy and time usage and is affordable. In another study, Karadag (2016) stated that materials heated with a microwave gave better results with a shorter process. Furthermore, according to research by Mihai (2016), the use of a drying machine can provide time effectiveness in increasing the productivity of the drying process to produce water hyacinth crafts.

Based on other research conducted by Radzuan (2020), on a kenaf leaf drying machine, data obtained for the drying process of kenaf leaves requires a temperature of 90–120°C. Given that the characteristics of gebang leaves and kenaf leaves are almost the same, but agel has a lower moisture content, we use this research as a reference for the temperature used in the drying machine oven, which is around 75–85°C.

## 2.2. Research Thinking Framework

In this framework of research thinking, the stages are explained starting from the analysis of the problems experienced by many Small Medium Enterprises (MSMEs) in the agel craft industry, namely the process of drying raw materials. This research was conducted from the development of its first version under the name Agel Dryer Machine. In the application of the first device, several shortcomings were still found related to machine use problems experienced by an SME and this became the background for further research. It is hoped that with the development of this second version of research, the resulting device will be appropriate and more perfect than the first version so that it can meet the needs of an SME in the agel craft

industry. The next stage is to develop alternative ideas for solving the problem in the form of an innovation, an appropriate technological machine for the agel drying process so that it can produce benefits for an SME. The stages of solving this problem include designing a new device by combining the concept of appropriate technology, a monitoring system and safety for the gas regulator section. Internet of Things (IoT) is developed and used in the system to make it easier for users to control and monitor devices via a smartphone (Setiawan et al., 2024). The framework of thinking used in the research is shown in Fig. 1.

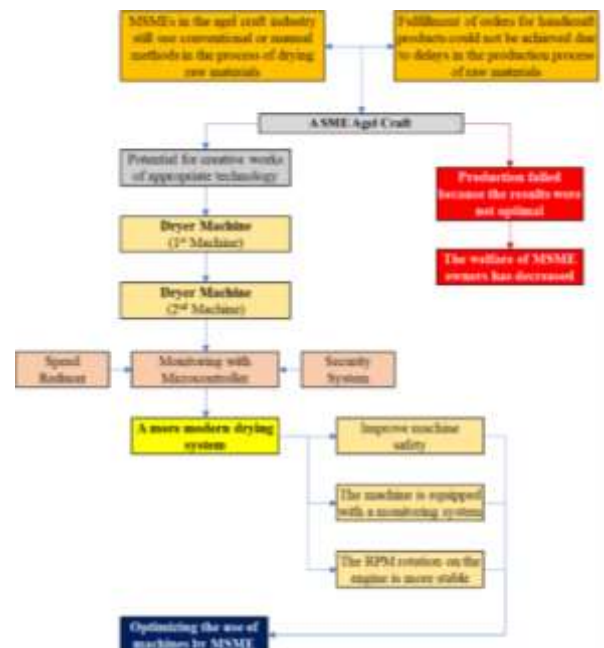


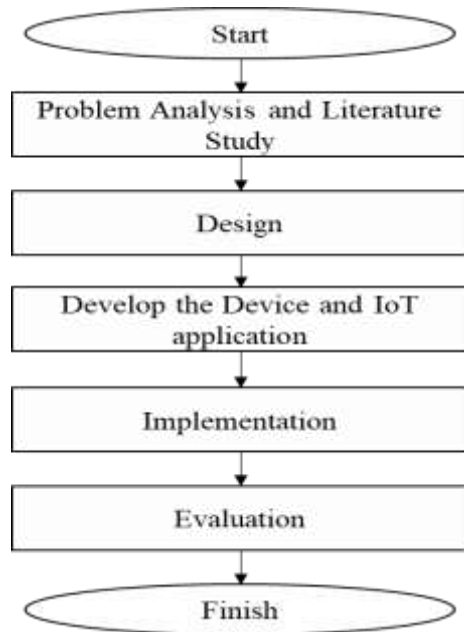
Fig. 1. Research thinking framework

## 3. Research Methodology

This study used the Research and Development (R&D) method, which is a research method to produce certain products and test the effectiveness of the products (Ma'ruf et al., 2024; Setiawan et al., 2021). This method is combined with the ADDIE model.

The machine development procedure applies to the ADDIE model which consists of five process stages including analysis, design, development, implementation, and evaluation (Salas-Rueda et al., 2020). The evaluation stage is used to determine the feasibility of the machine and the final stage of refinement (Chen & Setiawan, 2023). The initial stage is an analysis of the potential problems that are still being experienced from the application of the initial product. This ADDIE model research has several steps so that it becomes a machine that is ready for use. The flow of the

stages of the research process is conducted as shown in Fig. 2.



**Fig. 2.** Research and development (R&D) stages based on ADDIE model

## 4. Results and Discussion

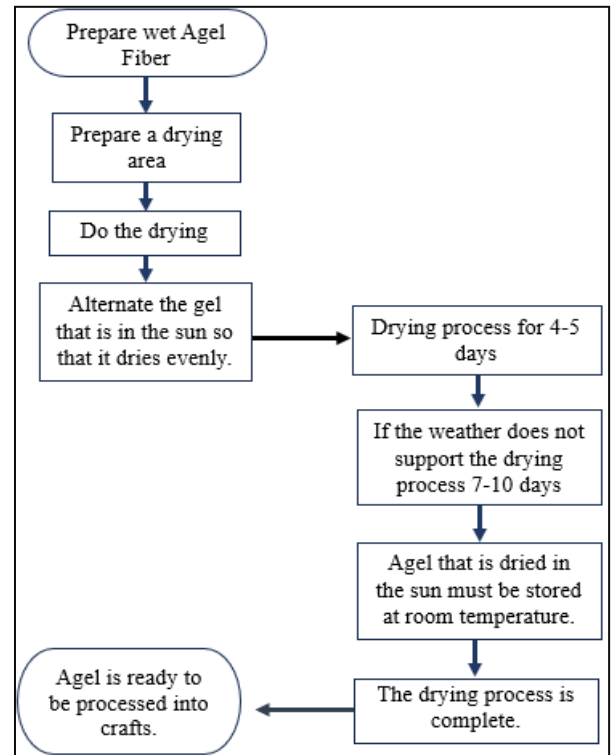
### 4.1. Drying Method Comparison

Currently, many SMEs continue to rely on conventional methods using sunlight for drying agel fiber. Unfortunately, this approach often leads to sub-optimal drying times, lengthy processes, and significant susceptibility to weather-related factors. As a result, our partners encounter a range of obstacles in the production of agel fiber craft materials.

In Fig. 3, there are drying steps using conventional methods. The main problem is during the drying process, this process takes 4 to 5 days. Additionally, additional time is required if weather conditions are not in clear condition.

To address this challenge, it is clear that the development of appropriate technological innovations is necessary. Fig. 4 shows a drying machine that has been developed.

Fig. 5 shows the framework of the drying step using a machine. The steps involved are fewer than those in the conventional method.



**Fig. 3.** Drying stage with conventional method



**Fig. 4.** Implementation of a drying machine

### 4.2. Product Specifications

Table 1 describes the specifications for the dryer machine, which has dimensions of 150 cm × 80 cm × 100 cm with a maximum production capacity of 10 kg. The materials used include the type of material for the frame, namely mild steel elbow 3 × 3 cm and mild steel plate with a thickness of 1.2 mm as material for fiber drying tube and machine cover. The power source for the machine's motor uses 900 VA home electricity. The source of the drying heat used is from an LPG stove with a temperature of 75–85°C. Based on testing on the drying machine that the researchers conducted

with the 10 kg gebang leaves using the temperature of 75–85°C for 4 hours, the gebang leaves were dried and not burnt as much as 7.5 kg.

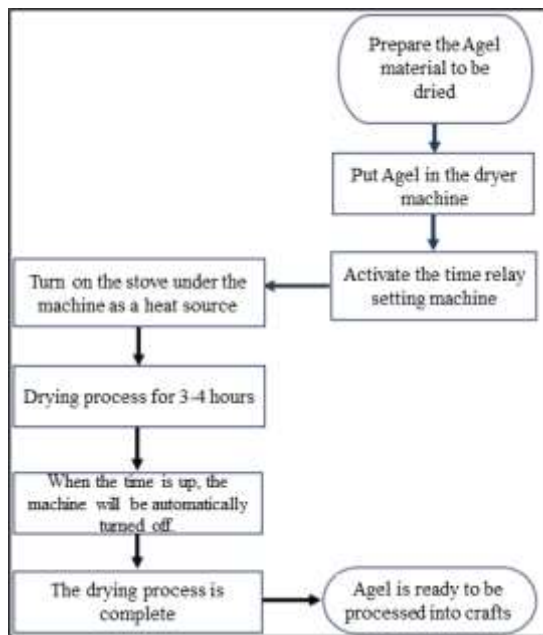


Fig. 5. Drying stage with machine

Table 1. Dryer machine specifications

No	Specification type	Information
1	Machine size	150 cm × 80 cm × 100 cm
2	Machine capacity	10 kg
3	Machine frame material	Mild steel elbow 3 × 3 cm
4	Drying tube material	Mild steel plate 1.2 mm
5	Drive mechanic	Electric motor ½ Hp
6	Heat source	LPG Stove
7	Monitoring system	IoT and application
8	Motor RPM	23.3 RPM
9	Power source for motor	Grid electricity 900 VA

#### 4.3. IoT System and Monitoring on Drying Machine

##### 4.3.1. IoT Implementation on Drying Machine

Apart from using an analog thermometer as a conventional temperature meter. The drying machine also uses a temperature sensor supported by the IoT to be able to monitor temperature with a smartphone. The

dryer monitoring system uses a temperature sensor placed inside the processing tube to detect the temperature of the air.

The circuit needs a power supply of 12 V DC electricity, stabilized with a step-down of 9 V, for the system to function. This ensures that the microcontroller receives the required power supply of 9 V. Once the power supply is on, the microcontroller and temperature sensor are activated. The circuit then looks for connectivity with the nearest Wi-Fi module that has been previously programmed. The connection between the circuit system and IoT is established through oven monitoring applications. Fig. 6 displays the IoT electronic circuit in the drying machine.

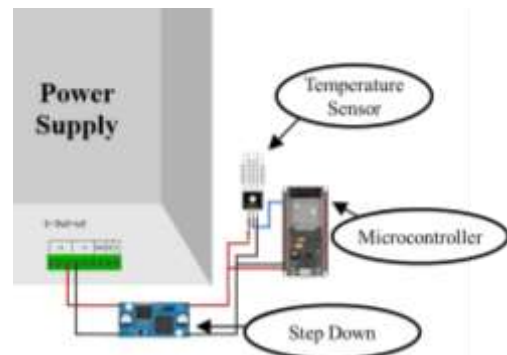


Fig. 6. IoT system electronic circuit

The temperature sensor will measure temperature values and convert them into digital signals. The digital signal from the temperature sensor will be read by the microcontroller. The microcontroller will transmit temperature and humidity data from the temperature sensor via the WiFi network to the cloud server. Temperature data that has been sent to the cloud will be stored in the server database in the application system.

Monitoring system administrators can view the current conditions of oven temperature and humidity via a web dashboard or smartphone application connected to the server. If the temperature or humidity of the oven exceeds the specified normal limits, the system can provide a warning in the form of a notification on the monitoring dashboard. Oven operators can immediately recognize abnormal conditions and take necessary action. By implementing this cloud-based IoT system, temperature and humidity in the oven can be remotely monitored without the need to physically

visit the oven location. The monitoring process becomes more efficient and effective.

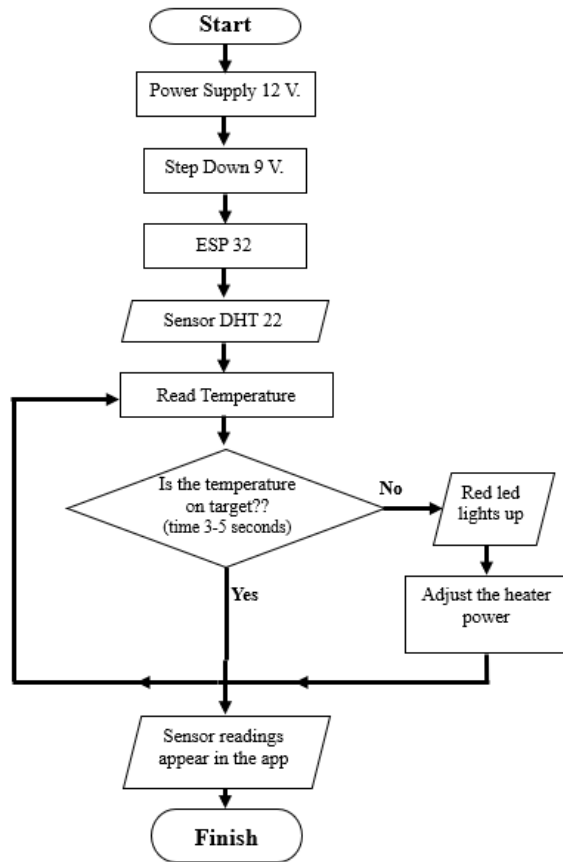


Fig. 7. Drying machine IoT working scheme

Fig. 7 explains the diagram of how the IoT system works and the use of the agel fiber drying machine, the initial stage of the machine's work starts from the process. The system is activated with a 12 V power source, the voltage is reduced to 9 V using a step-down converter to supply the microcontroller ESP 32, the temperature sensor DHT 22 is connected to the microcontroller to monitor the drying room temperature. The microcontroller reads temperature data from the temperature sensor periodically. If the temperature is detected to exceed the specified limit, the red LED will light up as a warning indicator. Temperature reading data from the sensor is displayed on an application connected to the microcontroller, enabling remote monitoring. This automation system allows precise temperature setting and control on the drying machine, increasing the efficiency of agel fiber production in an environmentally friendly way.

#### 4.3.2. Monitoring Applications

The application created from the kodular.io web is then extracted into Apk format so that it will become an application that can later be installed on the user's smartphone. The application display form can be seen in Fig. 8. This application has been programmed and connected to the sensor and microcontroller. Before entering the application monitoring system, users are asked to enter the username and password that they previously created, to ensure that application security and integration are more focused on connection.

At this stage, the application is only in the first stage of development, so a deficiency was found where the application had a delay of 2–3 seconds for sensor reading and data transmission. This factor was also caused by the type of microcontroller used. Research for monitoring systems has not yet reached the perfect phase so it needs to be further improved regarding the electronic components used. In this application there is a menu to periodically recap production results.

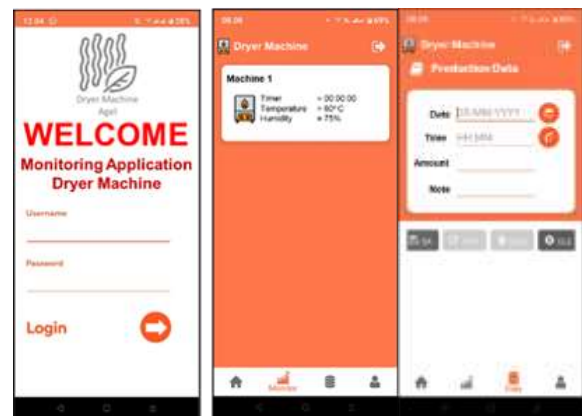


Fig. 8. Display of application monitoring dryer machine

#### 4.4. Dryer Machine Operation

The dryer machine works from the initial stage of inserting the wet gebang into the processing tub and then closing the machine until the final stage of the drying process is complete. Fig. 9 shows a drying machine turning wet gebang leaves into dried agel fiber.

The steps to use the Agel Dryer Machine are as follows:

- 1) The dryer machine works from the initial stage of inserting the wet gebang into the processing tank.



- 2) Then, close the machine processing tub, at this stage make sure the processing tub lock is completely closed.
- 3) The next stage is activating the machine and setting the time relay. At this stage, the monitoring process will also function, then the application is opened.
- 4) After the engine starts to rotate, turn on the engine heating stove.
- 5) Wait for the drying process for 2–2.5 hours.
- 6) The machine will turn off automatically according to the timer that has been previously set.
- 7) After the process is complete and the fiber has dried as shown in Fig. 11, take out the fiber and remove it from the processing tube.

Fig. 10 shows the first stage when the dryer machine tank is opened by the user.



**Fig. 9.** Overall view of the dryer machine



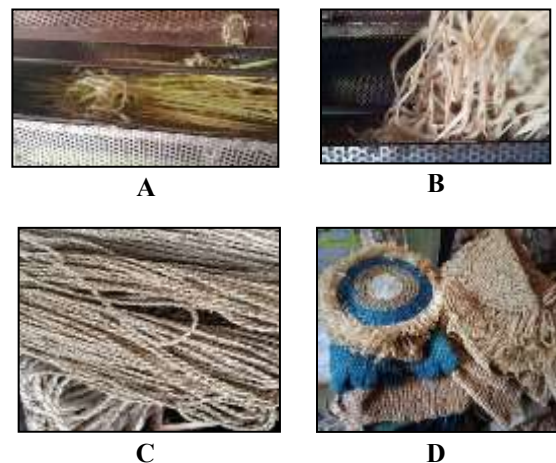
**Fig. 10.** Opening of the dryer machine tank



**Fig. 11.** Dryer machine tank with the dried fiber

#### 4.5. Implementation Result

The experiments using this machine were with capacities of 5 kg, 7.5 kg, and 10 kg. In this experiment, wet gebang leaves were able to dry for 2 hours, 3 hours, and 4 hours by maintaining the machine temperature between 75°C and 85°C. Fig. 12 presents the state of gebang leaves before and after the drying process to produce agel fiber, and the usage of the fiber to make handicrafts.



**Fig. 12.** Production of agel fiber from gebang leaves and handicraft making using the fiber. (A) Wet gebang leaves; (B) Dried gebang leaves; (C) Agel fiber from dried gebang leaves; (D) Agel fiber being made into handicrafts.

Fig. 12A–D show the images of processing wet gebang leaves into agel fiber and the making of handicrafts using the fiber. Fig. 12A shows the fresh and wet gebang leaves, which have a bright green color with a soft texture. It can also be seen that the leaves are still intact and have not dried out.

Fig. 12B shows gebang leaves that have been dried by the machine. The color of the leaves changes to brown after going through the drying process. The texture becomes stiffer and brittle compared to wet gebang leaves.

Fig. 12C demonstrates agel fiber from dried gebang leaves. In this image, there are results of agel fiber assembled from dried gebang leaves. This agel fiber has the appearance of a fine thread or rope and is brownish. This fiber is produced through a special processing process from dried gebang leaves.

Fig. 12D shows the usage of agel fiber for making handicrafts. The agel fiber is woven or knitted into various handicraft products, such as bags, hats, or other webbing. There are distinctive motifs and textures of agel fiber combined with the creativity of craftsmen.

Data related to machine testing, including time and dryer machine temperature, are presented in Table 2.

**Table 2. Dryer machine testing result**

Testing	Time (minutes)	Dryer machine temperature
1st Test (Capacity 5 kg wet gebang leaves)	0	23
	30	46
	60	63
	90	84
	120	83
2nd Test (Capacity 7.5 kg wet gebang leaves)	0	23
	30	46
	60	63
	90	83
	120	82
	150	81
3rd Test (Capacity of 10 kg wet gebang leaves)	180	84
	0	25
	30	49
	60	65
	90	84
	120	83
	150	84
	180	84
	210	84
	240	85

The final data of the research testing results reaches the research conclusions. It is known that the maximum weight of the wet gel used in the test is 10 kg. The drying temperature in the sun is between 25°C and 35°C or uncertain for 8 hours per day. Meanwhile, 10 kg of agel that is dried using a dryer machine is set at a temperature of 75–85°C and can dry in 4 hours. The water content of dried gebang leaves has an

average water content of 65%. Dried gebang leaves that have been processed by machine or called agel fiber have a dryness of 13–15%, with the water content in the fiber measured using the Amtast MC-7806 measuring instrument. The results related to the comparison can be seen in Table 3.

**Table 3. Dryer machine specifications**

No	Type of drying	Temperature	Amount	Drying time
1	Conventional method	25–35°C	10 kg (one drying)	40 hours (8 hours/day if sunny)
2	Dryer machine	75–85°C	10 kg	4 hours

#### 4.6. Result of Dried Agel Production

Calculations were conducted regarding the agel production after implementing the drying machine for 1 month of production to determine the increase in SME productivity before and after using this machine. This machine is capable of drying wet gebang leaves in one process with a maximum capacity of 10 kg, then craftsmen can get an average of 7.5 kg of dried agel fiber. Accumulatively, the conventional method is only able to produce 45 kg of dried agel fiber by one worker with 8 hours per day for one month (30 days) and under favorable weather conditions, whereas when using a dryer machine, it can produce 225 kg of dried agel fiber with continuous 24-hour operation, irrespective of weather condition.

#### 5. Conclusion

The study has successfully manufactured and implemented a drying machine that uses IoT system monitoring to address partner problems related to monitoring the agel fiber drying process. This system is equipped to monitor the drying process at features such as temperature, humidity, remaining time, and drying mark completion, and can function optimally according to the work method that has been designed. The dryer machine has dimensions of 150 cm × 80 cm × 100 cm and a production capacity of 10 kg. It features a 3 × 3 cm angle iron frame and a laser plate as part of its engine cover with a thickness of 1.2 mm. The machine uses 900 VA house electricity as its energy source for motor and an LPG gas stove with a temperature of 75–85°C as its drying heat source. With this machine, 10 kg of wet gebang leaves can be dried in 4 hours, producing 7.5 kg of dry agel, resulting in a



production of 225 kg per month. This machine creates a significant improvement compared to the conventional method that previously took 5–6 days to produce 10 kg of dried agel and was only able to produce 50 kg each month. By using this machine, the drying process becomes easier and the risk of MSDs due to inappropriate body posture while working can be reduced.

### Acknowledgments

The authors would like to thank PT. Engineer Tech Indonesia and CV. Inanri Sukses Bersama for supporting this work.

### References

- Albert, R., Sunardi, and Joko, S. (2022). Legal Protection of Biological Resources and Its Ecosystems in Indonesia. *Journal Social Science Research Network (SSRN)*, 1–16.
- Ali M.J, Faris M. A., and Anas A.J. (2021). Transient Behavior of Non-Toxic Natural and Hybrid Multi-Layer Desiccant Composite Materials for Water Extraction from Atmospheric Air. *Environmental Science and Pollution Research*, 28, 45609–45618.  
<https://link.springer.com/article/10.1007/s11356-021-13995-3>.
- Altshuller, G.S. (1984). *Creativity as an Exact Science: The Theory of the Solution of Inventive Problems*. Gordon and Breach Science Publishers.
- Arnau, S.S., Asensio, Cuesta., S.A., Seder, R. P. (2023). Musculoskeletal Disorders Risk Assessment Methods: A Scoping Review from A Sex Perspective. *Journal Ergonomics*, 1–20.  
<https://doi.org/10.1080/00140139.2023.2168767>.
- Benjamin, O.E., A., Stephen, A., Onyewuchi, E., and Uchenna, N. (2022). Design and Analysis of Energy and Exergy Performance of an LPG-Powered Fish Drying Machine. *Kejuruteraan*. 34(1), 117–129. [https://doi.org/10.17576/jkukm-2022-34\(1\)-11](https://doi.org/10.17576/jkukm-2022-34(1)-11).
- Bonfiglioli, R., Caraballo, A., Y., and Salmen, N. (2022). Epidemiology of Workrelated Musculoskeletal Disorders. *Current Opinion in Epidemiology and Public Health Journal*, 1(1), 18–24.
- Caponecchia, C., Coman, R.L., Gopaldasani, V., Mayland, E.C., and Campbell, L. (2020). Musculoskeletal Disorders in Aged Care Workers: A Systematic Review of Contributing Factors and Interventions. *International Journal of Nursing Studies*, 110, 103715.  
<http://dx.doi.org/10.1016/j.ijnurstu.2020.103715>.
- Chen, Y.T., & Setiawan, R.J. (2023). Energy Saving Solution for Welding Process: A SME Case Study of Fume Extractor. *2023 IEEE 6th International Conference on Knowledge Innovation and Invention (ICKII)*, Sapporo, Japan, 2023, 720–725.  
<https://doi.org/10.1109/ICKII58656.2023.10332684>.
- Edy L. (2023). Excess of Natural Resources Exploitation on Health and Environment by Foreign Investors. *Proceedings of the International Conference on Law, Economics, and Health (ICLEH), ASSEHR*, 723(1), 712–726.  
[https://doi.org/10.2991/978-2-38476-024-4\\_72](https://doi.org/10.2991/978-2-38476-024-4_72).
- Irwan, Juwita S., Deliyana I., and Tirsia K.M. (2023). Analysis Of Work Duration and Workload With Complaints Of Musculoskeletal Disorders (Msds) On Laundry Workers In Duingingi District The City Of Gorontalo. *International Journal of Health Science & Medical Research*, 2(1), 115–124. <https://doi.org/10.37905/ijhsmr.v2i2.13565>.
- Islam, S., & Mohammad, F. (2016). Sustainable Natural Fibres from Animals, Plants and Agroindustrial Wastes—An Overview. *Sustainable Fibres for Fashion Industry*, 31–44.  
[https://doi.org/10.1007/978-981-10-0566-4\\_3](https://doi.org/10.1007/978-981-10-0566-4_3).
- Karadag, A., Nazmiye, A., Kasapoğlu, K.N., Ozcelik, B. (2016). Effect of Microwave Technology on Some Quality Parameters and Aensory Attributes of Black Tea. *Czech Journal Food Sci*, 34(5), 397–405. <http://dx.doi.org/10.17221/5/2016-CJFS>.
- Karyadi, J., Hikam, F., Suganda, E., Muhklis, E., and Purwantana, B. (2019). Design of recirculated column dryer type for corn (*Zea mays* Linnaeus) drying. *IOP Conference Series: Earth and Environmental Science*, 365.  
<https://doi.org/10.1088/1755-1315/365/1/012034>.
- Kim, W.J., Park, H.J., and Jeong, B.Y. (2022). Article A Cross-Sectional Descriptive Study of Musculoskeletal Disorders (MSDs) of Male Shipbuilding Workers and Factors Associated the Neck. *Journal Applied Sciences*, 12, 3346.  
<https://doi.org/10.3390/app12073346>.
- Kumar, K., & Sekaran, A. (2014). Some natural fibers used in polymer composites and their extraction processes: A review. *Journal of Reinforced Plastics and Composites*, 33, 1879–

1892. <https://doi.org/10.1177/0731684414548612>.
- Lasano, N.F., Rahmat, A., Ramli, N.S., Bakar, M.F.A. (2018). Effect of Oven and Microwave Drying on Polyphenols Content and Antioxidant Capacity of Herbal Tea from *Strobilanthes crispus*. *Asian Journal of Pharmaceutical and Clinical Research*. 11 (6), 363–368. <http://dx.doi.org/10.22159/ajpcr.2018.v11i6.24660>.
- Ma'ruf, K., Setiawan, R. J., Darmono., Alam., A. A. K., Azizah, N., & Khosyati, N. E. (2024). Integration of Solar Photovoltaic System and Water Filter: A Sustainable Energy Solution for Clean Water Supply. *2024 International Conference on Electrical and Information Technology (IEIT)*, Malang. <https://doi.org/10.1109/IEIT64341.2024.10763095>.
- Maria W., Lusy & Ignatius R. (2022). Indonesian Kenaf Fiber Export Potential as A Raw Material For Various Industry. *International Journal of Economics, Business and Accounting Research (IJEBA)*, 6(1), 1980–1988.
- Mihai, B. (2016). How to use the DHT22 sensor for measuring temperature and humidity with the Arduino board. *ACTA Universitatis Cibiniensis*, 68 (1), 22–25. <http://dx.doi.org/10.1515/aucts-2016-0005>.
- Mochane, M.J., Mokheba, T.C., Mokhothu, T.H., Mtibe, Asanda, S., Ray E.R., Suprakas, S.I., and Daramola, I.D. (2019). Recent Progress on Natural Fiber Hybrid Composites for Advanced Applications: A Review. *CSIR Research Space*, 13, 159–198. <http://dx.doi.org/10.3144/expresspolymlett.2019.15>.
- Nikijuluw, V.P.H. (2017). Coastal Resources Conservation in Indonesia: Issues, Policies, and Future Directions. *Sumatra Journal of Disaster, Geography And Geography Education*, 1(1), 15–23. <https://doi.org/10.24036/SJDGGE.V1I1.31>.
- Nugroho, A., Abdurrohman, K., & Hestiawan, H. (2018, April). Effect of manufacturing method to tensile properties of hybrid composite reinforced by natural (agel leaf fiber) and glass fibers. In *Journal of Physics: Conference Series* (Vol. 1005, No. 1, p. 012004). IOP Publishing. <https://doi.org/10.1088/1742-6596/1005/1/012004>
- Nurhaliza, U.H., Centra, G. Anggoro, S.D., Oktaviandri, M. (2023). Design and Analysis Rolling Press Mechanism for producing Banana Stem as Natural Fiber. *Journal of Sustainable Mechanical Engineering*, 1, 8–11. <http://dx.doi.org/10.54378/josme.v1i1.5604>.
- Nuryanta, M.I., Aryaswara, L.D., Korsmik, R., Korsmik, O.K., Nugraha, A.D., Darmanto, S., Kusni, M., Muflikhun, M.A. (2023). The Interconnection of Carbon Active Addition on Mechanical Properties of Hybrid Agel/Glass Fiber-Reinforced Green Composite. *Journal Polymers*, 2411, 1–18. <https://doi.org/10.3390/polym15112411>.
- Radzuan, N.A.M., Tholibon, D., Sulong, A.B., Muhamad, N., and Haron, C.H.C. (2020). Effects of High-Temperature Exposure on the Mechanical Properties of Kenaf Composites. *Polymers*, 12(8), 1643. <https://doi.org/10.3390/polym12081643>.
- Salas-Rueda, R. A., Salas-Rueda, E. P., & Salas-Rueda, R. D. (2020). Analysis and design of the web game on descriptive statistics through the ADDIE model, data science and machine learning. *International Journal of Education in Mathematics, Science and Technology (IJEMST)*, 8(3), 245–260. <https://doi.org/10.46328/ijemst.v8i3.759>
- Saparudin, M.A., Setiawan, R.J., Budi, E., Puspito, A., and Fauzi, A. (2021). Design and Manufacture of Bamboo Handicraft Dryer Machine Based on LPG Gas. *Tadulako Science and Technology Journal* 2(1), 1–9. <https://doi.org/10.22487/sciencetech.v2i1.15555>
- Setiawan, R.J., Hidayat, H., Putra, P.P., Laksana, H.K., and Putra, P.I. (2021). Rancang Bangun Horizontal Gas Cutting Machine untuk Pemotongan Pipa. *Nozzle: Journal Mechanical Engineering*, 10(2), 34–44. <https://doi.org/10.30591/nozzle.v10i2.2572>.
- Setiawan, R.J., Ma'ruf, K., Suryanto., I.D., Azizah, N., Khosyati, N.E. (2024). Development of Rescue Drone integrated with Internet of Things Control for Beach Surveillance-Case Study at Glagah Beach. *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 603–608. <https://doi.org/10.1109/EECSI63442.2024.10776458>
- Widodo, Suparmin., Alfat K., and Basroni, A.N. (2020). Mapping Skill and Workplace Human Resources Leading Innovative Products Natural Fiber Vocational Village in Tuksono, Sentolo, Kulon Progo. *International Seminar Commemorating The 100th Anniversary of*

*Tamansiswa*, 1, 480–484.

You, W.S., and Jeong, B.Y. (2020) Work-related accidents and illnesses of older shipbuilding workers. *Journal of the Ergonomic Society Korea*, 39, 205–211.

<http://dx.doi.org/10.5143/JESK.2020.39.3.205>.

Zwawi, M. (2021). A Review on Natural Fiber Bio-Composites, Surface Modifications and Applications. *Journal Molecules*, 26, 1–28.  
<https://doi.org/10.3390/molecules26020404>.

## AUTHOR BIOGRAPHIES



**Khakam Ma'ruf.** He obtained his B. Ed in Mechanical Engineering from Yogyakarta State University. Now, he is a student in the Industrial Engineering study program, at Gadjah Mada University, Indonesia. He is active in research activity and experimental engineering.



**Rizal Justian Setiawan.** He obtained his A.Md.T and B.Ed. in Mechanical Engineering from Yogyakarta State University, M.M. in Marketing Management from Universitas Terbuka, and MSc in Industrial Engineering at Yuan Ze University. Currently, he is an active student in the M.A in College of Law and Politics at National Chung Hsing University and an entrepreneur in appropriate technology manufacturing.



**Darmono.** He is a lecturer at the Faculty of Engineering, Yogyakarta State University. He was appointed as a lecturer at the university with a scientific

concentration in Civil Engineering and Planning Education. He obtained his education degree such as Drs., Ir., and Dr. at Yogyakarta State University. Then, he obtained M.T. degree at Gadjah Mada University.



**Syukri Fathudin Achmad Widodo.** He is a lecturer in the Mechanical Engineering Education department at Yogyakarta State University, Indonesia. He possesses expertise in the field of mechanical engineering education and dedicates his/her career to teaching and conducting research in that area.



**Sumantri Sri Nugroho.** He is an experienced Technical Engineer working at the Appropriate Technology Development Center in the D.I. Yogyakarta. He possesses profound expertise in the field of engineering and applies his knowledge to the development of appropriate technology.



**Nur Evirda Khosyati.** She is an active student in the culinary technology education study program, culinary technology and fashion education department of Yogyakarta State University, Indonesia. She is active in research activity and experimental engineering.



**Nur Azizah.** She is a researcher from Sriwijaya University. She obtained S.Ked and M.D. degree at Sriwijaya University. Currently a student in the International Public Health (M.P.H) program at China Medical University, Taiwan (R.O.C).

# Brightness augmentation implementation to evaluate performance classification of masked facial expressions based on the CNN model

Desty Mustika Ramadhan<sup>1</sup>, Husni Mubarak<sup>1</sup>, Rianto<sup>1\*</sup>

<sup>1</sup>Department of Informatics, Siliwangi University, Tasikmalaya 46115, West Java, Indonesia

\* Corresponding author E-mail: rianto@unsil.ac.id

(Received 07 June 2024; Final version received 20 September 2024; Accepted 21 October 2024)

## Abstract

Deep learning methods with convolutional neural network (CNN) models have increasingly been applied to facial expression recognition. However, due to the recent pandemic, many individuals wear masks for work or health reasons, obstructing the complete visibility of their faces. This can impact social interactions, particularly in areas involving facial expression cues like the mouth. This study explores the application of CNNs in identifying facial expressions obscured by masks, focusing on the VGG16 and MobileNet architectures. Additionally, the research investigates the effects of data augmentation, including geometric and brightness augmentation, on the accuracy of facial expression classification. The findings indicate that the VGG16 architecture with cross-validation (VGG16-FLCV) outperforms MobileNet-FLCV in recognizing and classifying masked facial expressions. Data augmentation, particularly brightness augmentation, significantly enhances CNN model performance. For the VGG16-FLCV architecture, the brightness range (1.00, 1.25) yields the best accuracy, with a training accuracy of 81.73% and a validation accuracy of 70.71%. The most optimal brightness ranges for VGG16-FLCV are in the dark category (0.25, 0.50), (0.50, 0.75), and (0.75, 1.00), as well as the bright category (1.00, 1.25). Meanwhile, MobileNet-FLCV with brightness ranges (0.25, 0.50), (0.50, 0.75), (0.75, 1.00), (1.00, 1.25), and (1.25, 1.50) can be used as alternative brightness ranges without significant accuracy degradation. These findings provide valuable insights for improving the accuracy of masked facial expression recognition by applying appropriate data augmentation techniques.

**Keywords:** Brightness Augmentation, CNN, Cross-validation, Masked Facial Expressions

## 1. Introduction

Deep learning is a subset of machine learning. Therefore, it can be said that deep learning consists of a neural network with many layers and parameters. Most deep learning methods use a neural network (NN) architecture called deep NNs (Shinde & Shah, 2018).

The application of deep learning methods has gained prominence, particularly in facial expression recognition (FER). Most FER systems attempt to recognize expressions from a person's entire face. However, due to the pandemic in recent years, some people still wear masks to work or due to illness that requires them to wear a mask, which prevents their face from being fully visible (Castellano et al., 2021). The use of face masks has a negative effect. Psychologists report that it can confuse the reading of expressions, especially in the mouth region, which is very informative to help distinguish between expressions of sadness, disgust, fear, and surprise, thereby affecting social interactions (Yang et al., 2021).

There is research on classifying facial expressions with masks based on deep learning approaches. Grundmann et al. (2021) employed a strategy analytic approach using a multilevel regression model (logistic) to examine the effect of mask use on social judgment, investigate what facial cues are lacking when using masks to reduce expression recognition and explore how expression valence and associations related to the use of masks influence social judgment. Yang et al. (2021) proposed a facial expression type classification system for masked individuals based on a deep learning approach that applies convolutional neural network (CNN) models with MobileNetV2 and VGG19 architecture types. This research uses the M LFW-FER and M-KDDI-FER datasets, which have three types of facial expressions: positive, neutral, and negative. Hence, the masked facial expression classification process only considers facial expression classification with three expression categories. In addition, Castellano et al. (2021) used the CNN method with VGG16 and MobileNetV2 architecture types to recognize

emotions or expressions from the entire face and eye region of interest using the FER2013\_cropped dataset with seven types of expressions, namely angry, disgust, fear, happy, neutral, sad, and surprise. The study investigated how well the FER system can recognize expressions, even when individuals wear face masks and expressions are often confused with others when the face is covered.

In addition, data augmentation techniques in deep learning models are often used in the image classification process that can handle data scarcity. Data augmentation can increase the accuracy value of the trained CNN model because it provides additional data, enhancing variations in the dataset used (Waheed et al., 2020). The type of data augmentation commonly used in research is geometry transformation (Kandel et al., 2022). Pei et al. (2019) applied data augmentation techniques in face recognition using CNN to address the issue of insufficient data samples. The types of data augmentation used in their research are geometric transformation, brightness augmentation, image translation, image rotation, image zoom, and filter operation. The results showed that applying the CNN method with data augmentation can achieve an accuracy of 86.3% higher than the principal component analysis or the local binary pattern histogram method. In addition, according to Kandel et al. (2022), applying data augmentation techniques for classification on histopathology, especially on the Invasive Ductal Carcinoma Dataset. They used two types of data augmentation, geometric and brightness augmentation, applying eight brightness scales using four CNN models, namely Resnet50, DenseNet121, InceptionV3, and Xception. The results showed that the application of geometric augmentation provides better accuracy than the application of brightness augmentation. In addition, the CNN model provides better results without the application of data augmentation techniques. Hence, it can be hypothesized that in the research conducted by Kandel et al. (2022), the application of brightness augmentation significantly reduced model performance when extreme values were used.

Based on the entire description of the work, gaps still pave the way for future research, primarily related to the type of expression used in Yang et al.'s research, which only has three categories: positive, negative, and neutral. The system was designed using CNN architecture, as shown by the research of Castellano et al., which uses VGG16 and MobileNetV2 architecture, while Yang et al. used VGG19 and MobileNetV2. The choice of architecture in these studies becomes a reference for applying transfer learning to the VGG16 and MobileNet architectures in classifying masked facial expressions. In addition, Pei et al. applied data augmentation, especially brightness augmentation, solely to face recognition datasets. Kandel et al. only applied eight brightness scales to histopathology

datasets, particularly on the Invasive Ductal Carcinoma dataset.

Hence, this research is expected to provide a better understanding of masked facial expression recognition using deep learning techniques with the application of CNN architectures, namely VGG16 and MobileNet. In addition, data augmentation methods, such as geometric augmentation and brightness augmentation, will also be explored to evaluate their effect on the classification accuracy of masked facial expressions.

## 2. Related Work

Facial expression recognition systems mostly try to recognize expressions from a person's entire face. However, the pandemic has caused individuals to wear masks all the time, thereby causing their faces not to be fully visible. Currently, there is various research in classifying the types of facial expressions in masks. Table 1 shows the results of the related work and each proposed model's approach.

## 3. Methodology

This research methodology encompasses several key stages conducted systematically to achieve the research objectives. These stages include data collection, exploratory data analysis (EDA), data preparation, building the CNN model, model training, and model evaluation. The flow of these stages is illustrated in the flowchart shown in Fig. 1.

### 3.1. Data Collection

The research data from the Kaggle website and sample data for prediction tests were taken directly using a smartphone. The datasets used in this study include the MaskedDatasetFER, while 17 data samples were used for the prediction tests.

The explanation of the origin of the MaskedDatasetFER dataset is listed in the description on Kaggle (<https://bit.ly/3UP0oRz>), published in 2021. The MaskedDatasetFER dataset originated from the FER2013 dataset prepared by Pierre-Luc Carrier and Aaron Courville as part of an ongoing research project. A categorized dataset of masked individuals' facial expressions was created by artificially placing a face mask on the FER2013 dataset, forming a new dataset, MaskedDatasetFER.

MaskedDatasetFER consists of a dataset in the form of image files with image dimensions of  $48 \times 48$  pixels with a red, green, and blue (RGB) image type. The total data from the dataset is 20,484 image data, consisting of 15,531 training data and 4,953 validation data. Both data have seven types of expressions or class labels: angry, disgusted, fearful, happy, sad,

**Table 1.** Related work

Author	Model	Description	Advantages	Limitations
Yang et al. (2021)	CNN model by applying VGG19 and MobileNetV2 architecture types.	Focus on facial expression recognition (FER) of three types of expressions on the people in masks.	The model effectively enhances facial expression recognition accuracy by focusing on uncovered areas and surpasses other mask-aware recognition methods.	This model only classified three emotional categories: positive, neutral, and negative, so its generalization is still limited.
Castellano et al. (2021)	Implement CNN model with VGG16 and MobileNetV2 architecture types and apply ADAM optimizer type.	Focus on introducing automatic expression of the expression face when wearing a mask.	This research has succeeded in developing an effective system for recognizing facial expressions only from the eye area.	This system has limitations in managing negative emotions, which often confuse expressions of sadness with anger or fear.
Pei et al. (2019)	The CNN model is used by applying VGG16 architecture types and the cross-validation method.	Focus on recognition face, recognition through learning deep learning using data augmentation based on experiment orthogonal.	The data augmentation method successfully increased the accuracy of facial recognition to 98.1% for class attendance.	The data collection process and the required orthogonal experiments can be challenging to implement.
Kandel et al. (2022)	Using CNN mode by applying Resnet50, DenseNet121, InceptionV3, and Xception.	Focus on brightness as an augmentative technique for image classification on an AGY dataset of invasive ductal carcinoma dataset.	Applying geometric augmentation techniques is more effective than brightness in improving CNN performance.	Brightness augmentation, especially at extreme values, can degrade model performance and not improve classification results.
Grundmann et al. (2021)	The multilevel regression (logistic) method was used with the nlmnb optimizer, and the Glmer method was applied.	Focusing on face masks reduces expression recognition accuracy and perceived proximity.	This research reveals a significant impact of face mask use on the accuracy of emotion recognition and social judgment and provides essential insights for mask-related policy making.	Face masks have been shown to reduce the ability to classify emotional expressions accurately and may decrease feelings of closeness, particularly in older adults.
Cotter et al. (2020)	Implement CNN models with MobileNet, MobileEx, and ResNet architecture types.	Focus on introducing facial expression recognition on smartphones.	The proposed MobiExpressNet model has more than 5 times smaller size and FLOPs than the smallest MobileNet model, with 67.96% accuracy on the FER2013 dataset, making it very attractive for real-time smartphone applications.	The performance and accuracy of the ExpressNet Mobile model have not been tested in real-world conditions on smartphone devices, which needs to be considered in further development.
Genc et al. (2020)	Using the Wizard of Oz	Focus on face mask design to reduce occlusion of facial expressions.	Electrochromic technology in smart masks enhances communication by displaying facial expressions.	The study was limited to a small sample and did not test the automated mechanism for real-world use.
Merghani et al. (2020)	Implement the FME algorithm and apply SMO, CASMEII, and SAMM methods.	Focus on creating a new adaptive mask for region-based facial micro-expression recognition by defining 14 new rois based on the most frequently used action units (au).	Region-based and adaptive mask methods for recognizing facial micro-expressions show promising accuracy, with competitive results compared to deep learning approaches on the SAMM dataset.	The accuracy of this method is still relatively low compared to other methods, and this research only evaluated two datasets without considering the potential combination with deep learning approaches in the future.

surprised, and neutral. The sample data from MaskedDatasetFER is provided in Fig. 2.

Meanwhile, the data samples used for the prediction tests have 17 image data samples with dimensions of  $300 \times 300$  and are RGB image types. The sample data for the prediction test is shown in Fig. 3.

### 3.2. Exploratory Data Analysis

Exploratory data analysis (EDA) is a process for exploring the dataset to be used. In this research, EDA includes two main aspects: data understanding and data visualization.

**Table 1. Cont'd**

<b>Author</b>	<b>Model</b>	<b>Description</b>	<b>Advantages</b>	<b>Limitations</b>
Yang et al. (2022)	Cross-attention-based and vision transformer models with CNN architecture types were used, including VGG19, MobileNet1, ResNet, and ViT, and the application of RAN, ACNN, and OADN methods.	Focus on Facial Expression Recognition based on Face Parsing and Vision Transformer.	The proposed method combines Transformer face parsing and vision models with a cross-attention mechanism to improve facial expression recognition accuracy with masks, surpassing the existing FER method.	This research has not tested the method in real-world contexts or broader scenarios beyond the dataset used.
Agrawal et al. (2020)	Using the CNN model and the FER2013 dataset.	Focus on studying the effect of kernel size and number of filters on facial expression recognition accuracy.	Presenting two new simple and effective CNN architectures, achieving 65% accuracy on the FER-2013 dataset.	This research is limited to the FER-2013 dataset, which may not be generalizable to other datasets or real-world applications.
Ding et al. (2020)	The CNN model was used with Resnet50, VGG16, and AffecNet architecture, and the OADN approach was applied.	Focus on occlusion-adaptive deep network for robust facial expression recognition.	This method improves the accuracy of expression recognition by handling occluded facial features and dividing the feature map into independent facial blocks for better robustness to occlusion.	This research requires high computational resources for model training and application. It may be less than optimal for very subtle or external expression variations not covered by the dataset.
Cheng et al. (2019)	Using model FSNet.	Focus on enhanced face segment or deep feature learning for face recognition.	FSNet enhances identity discrimination by exploiting local facial features through semantic segmentation and parsing maps while integrating global and local information.	The method may struggle with significant intra-personal variations or extreme facial condition changes not fully covered by the parsing maps.
Li et al. (2020)	Implementing the CNN model with VGG16 architecture and applying attention mechanisms and LBP approaches.	Focus on CNN-based attention mechanisms for facial expression recognition.	The proposed network integrates LBP features and an attention mechanism to enhance performance in facial expression recognition, demonstrating superior results on multiple datasets, including a newly collected one.	The method is currently limited to 2D images and does not address video data, 3D face datasets, or depth images, which could restrict its applicability.
Farkhod et al. (2022)	Using CNN model with Haar-Cascade classifier.	Focus on developing real-time landmark-based emotion recognition CNN for masked faces.	The proposed method achieves a high accuracy of 91.2% in image-based emotion recognition for masked faces by utilizing landmark-based features and a CNN model, showing effective performance compared to existing models.	Real-time emotion detection accuracy is lower due to biases and noise, such as image blurriness and poor lighting, and the method requires further development to improve performance under such conditions.

### 3.2.1. Data Understanding

The data understanding process in this research involves several crucial steps. Firstly, understanding the number of datasets in MaskedDatasetFER. Next, identify the types of classes present in the dataset. The following step involves evaluating the amount of data within each class in the dataset and understanding the data shape to ensure that the input dimensions are appropriate when building the CNN model. Since this research uses image data, verifying the RGB pixel values is necessary to ensure that the data used is indeed RGB image data.

### 3.2.2. Data Visualization

The data visualization process in this study involves two steps, including displaying the

distribution of the MaskedDatasetFER dataset for both training and validation data and displaying images found in each class within the dataset.

### 3.3. Data Preparation

The data preparation stage is a crucial process involving the preprocessing of the dataset before the model training phase. In this research, the dataset was divided into two parts: training data and validation data. The data was split using various ratios, including 50:50, 60:40, 70:30, 80:20, and 90:10, to evaluate the model's performance based on different training and validation data proportions. Additionally, data augmentation techniques were applied, including brightness adjustment, rescaling, zooming, rotation, and other transformations to enhance the variety and quality of the training data. The brightness



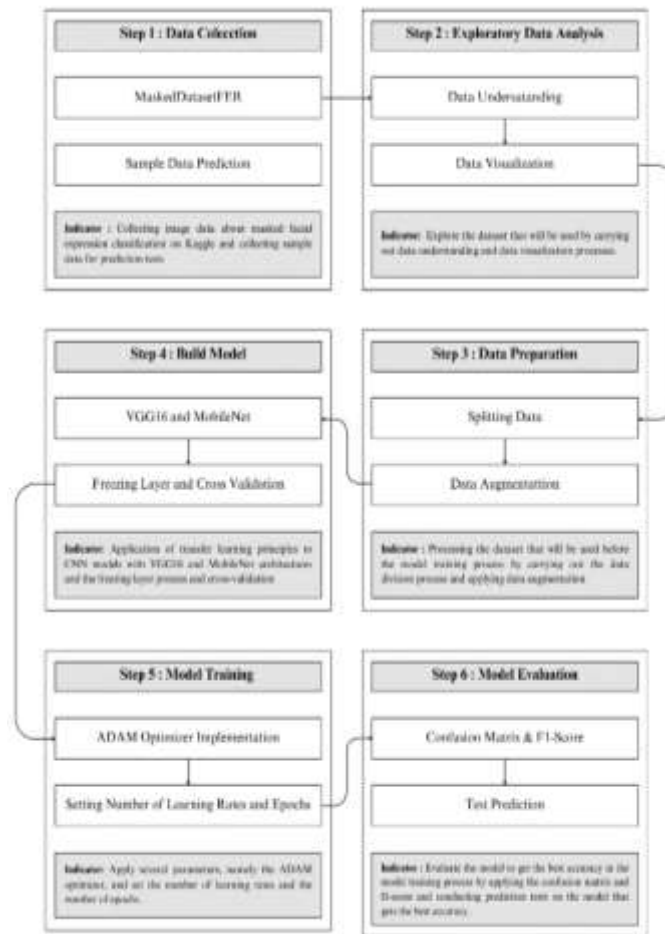


Fig. 1. Research flowchart.



Fig. 2. Sample MaskedDatasetFER.



Fig. 3. Sample prediction test.

augmentation specifically employed eight parameters, as used in the research by Kandel et al. (2022), with brightness ranges of 0.00–0.25, 0.25–0.50, 0.50–0.75, 0.75–1.00, 1.00–1.25, 1.25–1.50, 1.50–1.75, and 1.75–2.00. The results of these data augmentation techniques are illustrated in Fig. 4.

### 3.4. Building the CNN Model

Developing the CNN model in this research involves applying transfer learning to pre-trained architectures, namely VGG16 and MobileNet. The

CNN architectures used in this study are VGG16-FL and MobileNet-FL, which apply the freezing layer technique to specific layers and modify the top layers. The detailed architectures of VGG16-FL and MobileNet-FL are illustrated in Fig. 5 and Fig. 6.

### 3.5. Model Training

At this stage, the CNN model that has been constructed is trained to achieve optimal performance. The training process involves the application of



Fig. 4. Data augmentation.

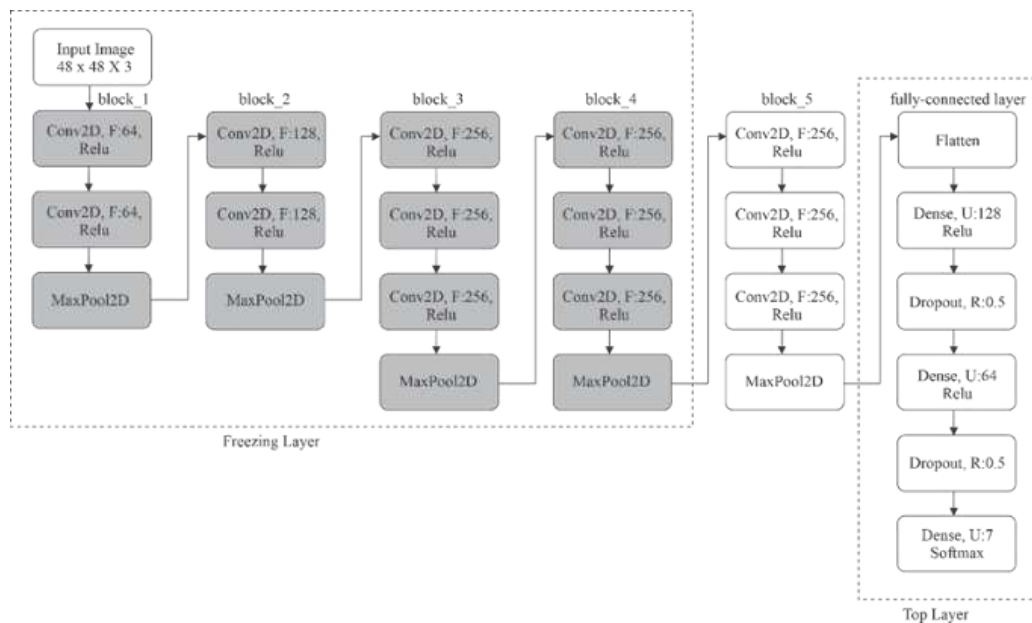


Fig. 5. VGG16-FL architecture.

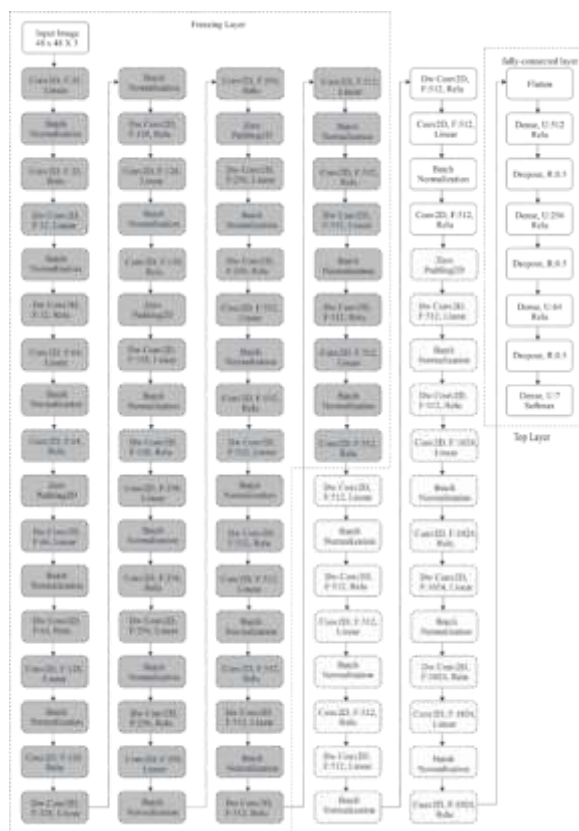
various hyperparameters, including the ADAM optimizer. In this process, fine-tuning is performed by setting the learning rate at three levels: 0.001, 0.0001, and 0.00001. The number of epochs is also set to 100 to ensure the model can learn effectively from the available data.

The model training process applies the cross-validation method with k-fold values ranging from 2 to 10. This approach evaluates and improves the model's ability to generalize to the new data and address the imbalanced data issue. By comprehensively assessing the model's performance across various subsets of data, the model is expected to deliver consistent and accurate results, mainly when tested with uneven class distribution data.

$$Recall = \frac{(TP)}{(TP+FN)} \quad (3)$$

$$F1\ Score = \frac{(2 \times Recall \times Precision)}{(Recall + Precision)} \quad (4)$$

T represents true, P represents positive, N represents negative, and F represents false.



#### 4.1. Analysis Model CNN

Overfitting and underfitting conditions are two of the problems that cause inaccurate and suboptimal prediction results. Overfitting conditions can occur when a (NN that is overly dependent on the training set learns incorrect mappings that work well in the training set but perform poorly in the validation or testing set (Zhang et al., 2019). In addition, models trained with an unbalanced data set may become overfitted to training samples from underrepresented, resulting in poor generalization during test time (Li et al., 2021). There are several alternatives to handling models that experience overfitting conditions used in this study, including adding the application of data augmentation, which has been empirically proven to reduce overfitting with very high dimensional data by increasing the amount and variety of training data (Rice et al., 2020). Another way is the application of dropout by randomly discarding information targeting each hidden node of the NN during the training phase (Choe et al., 2019). A possible factor causing underfitting is that the NN architecture is too simple and has too few hidden layers or trainable parameters, making it not powerful enough to capture complex data characteristics (Zhang et al., 2019).

**Fig. 6.** MobileNet-FL architecture.

The system evaluation stage, which involves assessing the accuracy achieved in the research, calculates precision and recall values using the confusion matrix and F1-score methods. The calculations for accuracy, precision, recall, and F1-score based on the confusion matrix are presented in equations (1), (2), (3), and (4), according to Castellano et al. (2021).

values and experience overfitting conditions. Data augmentation and dropout are applied to overcome the overfitting conditions, but in experiments without cross-validation, it is not effective enough to address overfitting. Therefore, the cross-validation method is applied to reduce overfitting in this research. In addition, overfitting can occur due to several factors, such as the dataset used having an imbalance of data between class labels for both training data and validation data. Cross-validation is one of the most widely used data resampling methods to estimate the true prediction error of a model and one of the methods used to prevent overfitting conditions (Berrar et al., 2018).

The application of the cross-validation method on both VGG16-FL and MobileNet-FL architectures in this study proved to be quite effective in handling overfitting. This is shown by the accuracy value obtained, which increased significantly from the experiment without cross-validation. The application of the number of k-folds in the cross-validation method also affects the accuracy value obtained, namely in the VGG16-FL cross-validation (VGG16-FLCV) architecture, resulting in the highest accuracy value at a value of 7-fold cross-validation. Meanwhile, in the MobileNet-FL cross-validation (MobileNet-FLCV) architecture, the highest accuracy value is at the 8-fold cross-validation value. Subsequently, brightness augmentation is applied to measure the effect on the accuracy value obtained. The graphs showing the accuracy and loss values on the VGG16 and MobileNet architectures and the application of cross-validation and brightness augmentation methods are illustrated in Fig. 7 to Fig. 14.

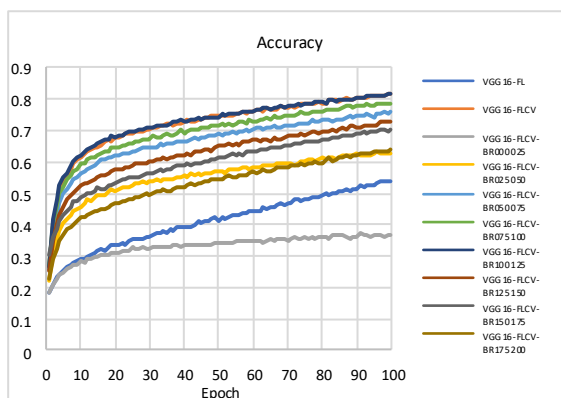


Fig. 7. Training accuracy with VGG16-FL.

Fig. 7 shows that the VGG16 architecture applying the cross-validation method has a higher training accuracy (VGG16-FLCV) than the VGG16 architecture without the cross-validation method. In addition, the graph shows that the application of brightness augmentation can increase the training accuracy value obtained precisely in the brightness range (1.00, 1.25), which obtained a training accuracy

value of 81.73%. Meanwhile, the application of the brightness range (0.00, 0.25) has the lowest accuracy value compared to other experiments, with a training accuracy value of 36.78%. Hence, it can be concluded that applying the brightness range (0.00, 0.25) can significantly reduce the training accuracy value. The comparison graph of the loss value between models that do not apply the cross-validation method with models that apply the cross-validation method and the comparison of each application of brightness augmentation on the VGG16-FLCV architecture is illustrated in Fig. 8.

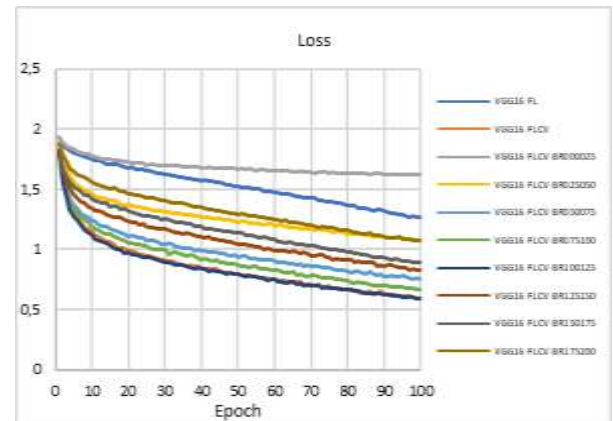


Fig. 8. Training loss with VGG16-FL.

Meanwhile, the training loss value graph in Fig. 8 shows that the smallest training loss value is found in the VGG16 architecture that applies the cross-validation method and brightness augmentation with the brightness range (1.00, 1.25). The highest training loss value listed in the graph is on the architecture applying brightness range (0.00, 0.25). In addition, the validation accuracy and loss graphs with the VGG16 architecture can also be seen in Fig. 9 and Fig. 10.

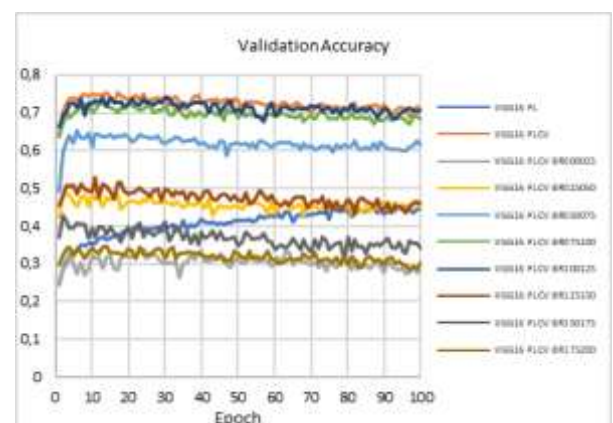


Fig. 9. Validation accuracy with VGG16-FL.

Fig. 9 shows a comparison graph of validation accuracy illustrating differences in the previous training accuracy graph. This validation accuracy shows relatively stable results, with only minor



fluctuations in the form of increases or decreases in validation accuracy values. However, the highest validation accuracy value is observed in experiments applying the cross-validation method (VGG16-FLCV) and in experiments applying brightness augmentation with brightness range (1.00, 1.25). The validation accuracy value in the VGG16-FLCV experiment is 70.61% and 70.71% in the brightness range (1.00, 1.25) experiment. In addition, the validation accuracy value that has the lowest value is in the experiment applying brightness augmentation (0.00, 0.25) with the resulting validation accuracy value of 28.26%.

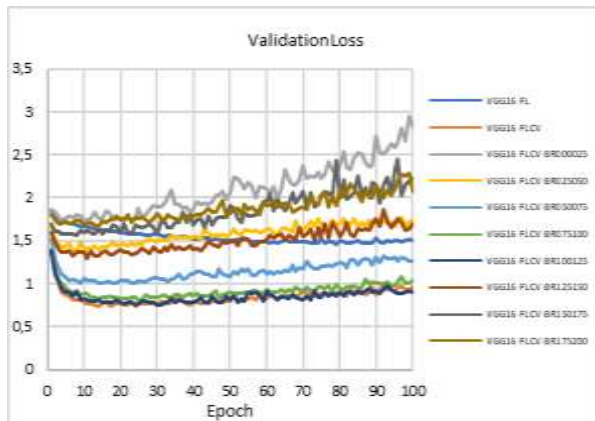


Fig. 10. Validation loss with VGG16-FL.

The graph in Fig. 10 shows that the experiment that applies the cross-validation method and the experiment that applies the brightness range (1.00, 1.25) have relatively low validation loss values compared to the validation loss value in other experiments. The experiment applying the brightness range (0.00, 0.25) has the highest validation loss value compared to other experiments.

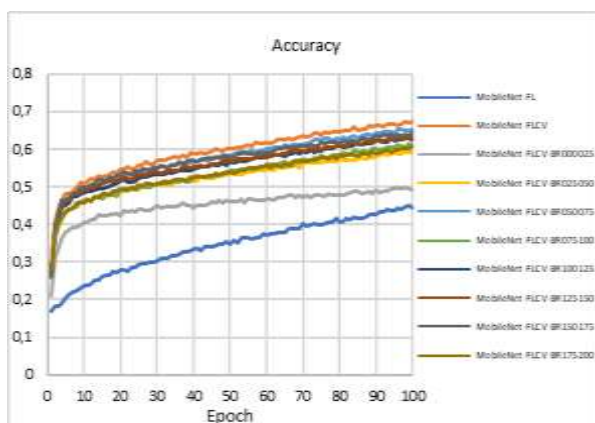


Fig. 11. Training accuracy with MobileNet-FL.

Based on the graph in Fig. 11, the MobileNet architecture applying the cross-validation method (MobileNet-FLCV) has a higher training accuracy compared to the MobileNet architecture without the cross-validation method. In addition, the graph shows

that the application of brightness augmentation with the MobileNet architecture applying the brightness range (0.00, 0.25) can significantly reduce the accuracy value. In addition, the training accuracy value with the highest value is the experiment applying the cross-validation method without employing brightness augmentation, obtaining a training accuracy value of 67.47%. The lowest training accuracy value was found in the MobileNet architecture experiment without cross-validation, resulting in a training accuracy value of 44.46%.

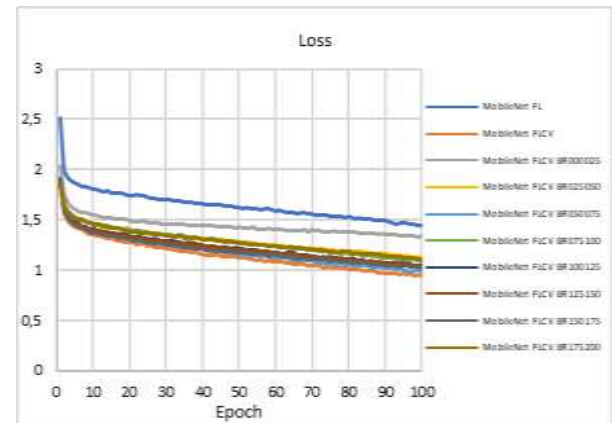


Fig. 12. Training loss with MobileNet-FL.

Meanwhile, the training loss value graph in Fig. 12 shows that the lowest training loss value is the MobileNet architecture applying the cross-validation method without brightness augmentation, and the highest training loss value is the experiment without applying the cross-validation method.

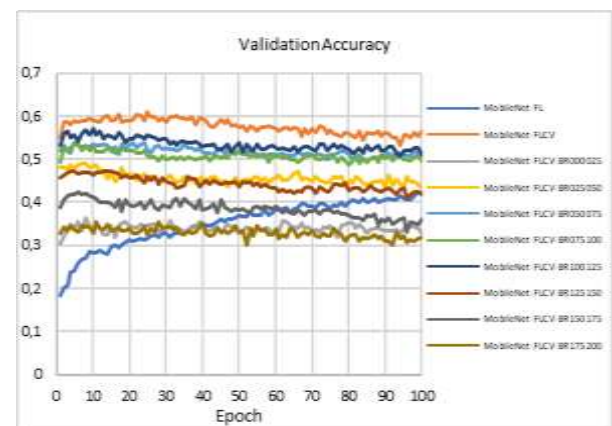


Fig. 13. Validation accuracy with MobileNet-FL.

Fig. 13 shows a comparison graph of validation accuracy showing differences in the previous training accuracy graph. This validation accuracy graph shows relatively constant results, indicating similarities in the experiments using the VGG16 architecture. Hence, it can be said that the changes that occur in the graph, both in the form of an increase or decrease in the validation accuracy value that occurs, are not too

significant. However, the highest validation accuracy value is the experiment that applies the cross-validation method or MobileNet-FLCV with a validation accuracy value of 56.37%. Meanwhile, the lowest validation accuracy value is found in the experiment applying the brightness range (1.75, 2.00) with a validation accuracy value of 31.80%.

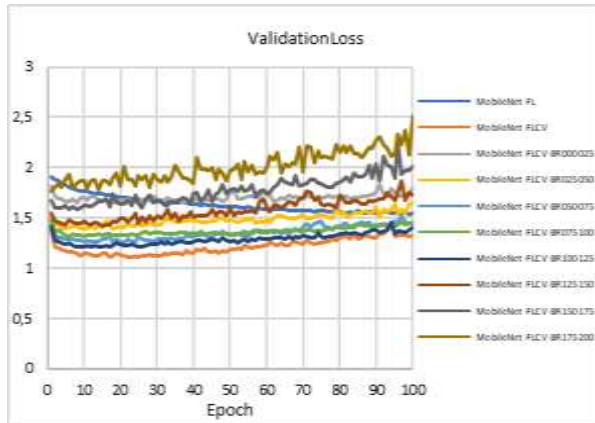


Fig. 14. Validation Loss with MobileNet-FL.

The graph in Fig. 14 shows that the experiment that applies the cross-validation method has a relatively small validation loss value compared to the validation loss value in other experiments, and the highest validation loss value is in the experiment with the brightness range (1.75, 2.00).

Based on the graphical results in Fig. 7 to Fig. 14, it can be said that the application of brightness augmentation can reduce and increase the accuracy values obtained. The application of brightness augmentation in the VGG16-FLCV architecture that can be applied in the model training process on the classification of masked facial expressions is the brightness range (0.25, 0.50), (0.50, 0.75), (0.75, 1.00), and (1.00, 0.25) ranges do not experience a very significant decrease in accuracy value. The application of brightness augmentation with a brightness range (1.00, 0.25) can increase the accuracy value from the accuracy value results that do not apply brightness augmentation in the VGG16-FLCV architecture. The accuracy values obtained from the application of brightness range (1.00, 0.25) with VGG16-FLCV architecture are 81.73% (training accuracy) and 70.71% (validation accuracy). Meanwhile, the model evaluation process is also carried out on the VGG16-FLCV architecture that applies brightness augmentation. Application of brightness range (0.00, 0.25) produces lower values compared to other experiments, namely with a precision value of 36.35%, recall of 34.03%, and F1-score of 34.81%. In addition, the application of brightness augmentation produces the highest value of prediction accuracy, precision, recall, and F1-score, namely in the use of brightness range (0.75, 1.00) in 7-fold cross-validation with a value of precision 74.51%, recall 72.38%, and F1-

score 73.22%, and the use of brightness range (1.00, 1.25) in 7-fold cross-validation with precision 76.23%, recall 74.16%, and F1-score 74.97%.

Meanwhile, the application of brightness augmentation in the MobileNet-FLCV architecture mostly decreases the accuracy value of the experiment that does not apply brightness augmentation. The application of brightness augmentation that has a fairly high accuracy value among other brightness augmentation applications is the brightness range (1.00, 1.25), which obtained a training accuracy value of 62.66% and a validation accuracy value of 51.21%. So, it can be said that the application of brightness augmentation in the MobileNet-FLCV architecture that can be applied in the model training process on the classification of masked facial expressions is in the brightness range (0.25, 0.50), brightness range (0.50, 0.75), brightness range (0.75, 1.00), brightness range (1.00, 0.25) and brightness range (1.25, 1.50). Meanwhile, the model evaluation results with the application of brightness augmentation and cross-validation show that the Mobilenet architecture can have smaller precision, recall, and F1-score values compared to the application of the VGG16 architecture. The precision, recall, and f1-score values are quite high compared to the other experiments in the application of MobileNet-FLCV architecture found in the experiment without applying brightness augmentation with 8-fold cross-validation, which is 63.58% precision, 60.93% recall, and 61.26% f1-score.

This study compares several models based on five performance metrics: precision, recall, F1-score, accuracy, and validation accuracy. The models include various VGG16 and MobileNet variants, each evaluated under different configurations, as shown in Table 2.

The results indicate that the application of cross-validation significantly improves the performance of both VGG16 and MobileNet models. However, the effects of brightness augmentation show variability, suggesting that careful adjustment of augmentation parameters is essential to optimize model performance.

Based on research conducted by Kandel et al. (2022), the application of geometric augmentation provides a better accuracy value compared to the application of brightness augmentation. In addition, the CNN model, without applying data augmentation techniques, gives better results than the application of brightness augmentation. However, in this study, the application of brightness augmentation can increase the resulting accuracy value, namely in the application of VGG16-FLCV architecture with the brightness range (1.00, 1.25) in the process of masked facial expression classification. This research proposes several brightness range parameters that can be used for the masked facial expression classification process for both darkness and brightness categories in each architecture used, namely VGG16-FLCV and MobileNet-FLCV. In addition, there are similarities



**Table 2.** Comparison of value precision, recall, f1-score, accuracy, and validation accuracy.

Techniques used	Precision	Recall	F1-Score	Accuracy	Validation accuracy
VGG16-FL	0.1766	0.1689	0.1713	0.5413	0.4454
VGG16-FLCV	0.7716	0.7535	0.7617	0.8165	0.7061
VGG16-FLCV-BR000025	0.3635	0.3403	0.3481	0.3678	0.2826
VGG16-FLCV-BR025050	0.5272	0.4835	0.4950	0.6303	0.4491
VGG16-FLCV-BR050075	0.6777	0.6527	0.6604	0.7592	0.612
VGG16-FLCV-BR075100	0.7451	0.7238	0.7322	0.7884	0.682
VGG16-FLCV-BR100125	0.7623	0.7416	0.7497	0.8173	0.7071
VGG16-FLCV-BR125150	0.5496	0.5280	0.5344	0.7270	0.4573
VGG16-FLCV-BR150175	0.5395	0.4289	0.4571	0.7040	0.3401
VGG16-FLCV-BR175200	0.4211	0.3468	0.3608	0.6372	0.3028
MobileNet-FL	0.2169	0.1723	0.1843	0.4446	0.4191
MobileNet -FLCV	0.6586	0.6351	0.6440	0.6875	0.5512
MobileNet-FLCV-BR000025	0.4062	0.3941	0.3959	0.5214	0.3602
MobileNet-FLCV-BR025050	0.5107	0.4968	0.4966	0.6053	0.4437
MobileNet-FLCV-BR050075	0.6117	0.5843	0.5936	0.6699	0.5094
MobileNet-FLCV-BR075100	0.5992	0.5500	0.5655	0.6102	0.4941
MobileNet-FLCV-BR100125	0.5961	0.5746	0.5798	0.6443	0.5230
MobileNet-FLCV-BR125150	0.5824	0.4636	0.4887	0.6298	0.4199
MobileNet-FLCV-BR150175	0.4908	0.4128	0.4253	0.6432	0.3734
MobileNet-FLCV-BR175200	0.5021	0.3675	0.3926	0.6089	0.3102

between masked facial expression classification research contained in Yang et al.'s (2021) research and Castellano et al.'s (2021) research in terms of using CNN architecture, which is a reference in this research. However, there are differences in this study, namely, in this study, there is an additional method to optimize the accuracy value obtained by adding the cross-validation method to the VGG16 and MobileNet architectures.

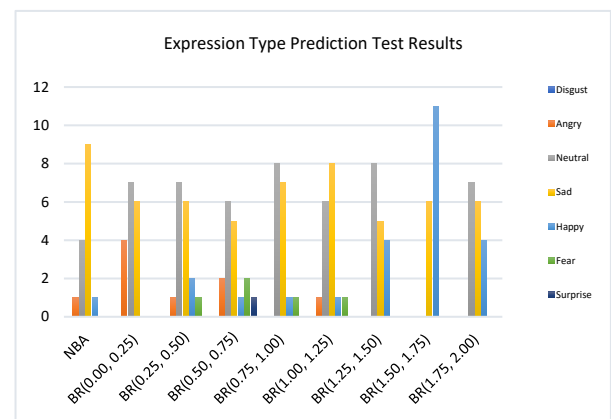
#### 4.2. Analysis Prediction Test

The prediction test process is carried out using prediction sample data to test whether the model can predict masked facial expressions with real data. The prediction test process is carried out using a model with the application of the VGG16-FLCV architecture in 7-fold cross-validation and MobileNet-FLCV 8-fold cross-validation and with the application of brightness augmentation. The results of the prediction test process can be seen in Fig. 15 to Fig. 18.

Fig. 15 shows the results of the types of expressions produced in the 17 samples used in the prediction test process on the VGG16-FLCV architecture based on the techniques used. In addition, based on this figure, it shows that the model that can predict the most types of expressions is the model that

applies the brightness range (0.50, 0.75), which in 17 samples predicts angry, neutral sad, happy, fear, and surprise expressions. Meanwhile, the model that can predict the least types of expressions is the model that applies the brightness range (1.50, 1.75).

The prediction test was carried out on 17 data samples; each sample has a difference in predicting the type of facial expression in the application of the VGG16-FLCV architecture. The graph that shows the results of the expression type prediction test for each sample can be seen in Fig. 16.

**Fig. 15.** Expression type prediction test results VGG16-FLCV.

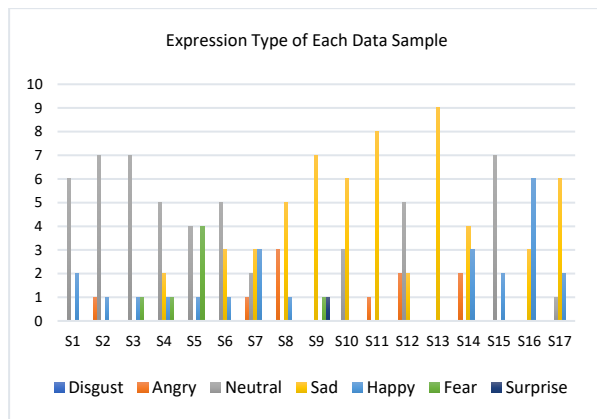


Fig. 16. Expression type of each data sample VGG16-FLCV.

Fig. 16 shows the number of facial expression prediction results for each data sample used in the prediction test with the VGG16-FLCV architecture. Data samples that have the highest number of expression predictions are the fourth (S4) and seventh (S7) data samples. The fourth data sample (S4) is predicted to have neutral, sad, happy, and fear expression types. Meanwhile, the data sample that has the least number of predictions, with one type of expression, is the 13<sup>th</sup> data sample (S13), which is predicted to have a sad expression type.

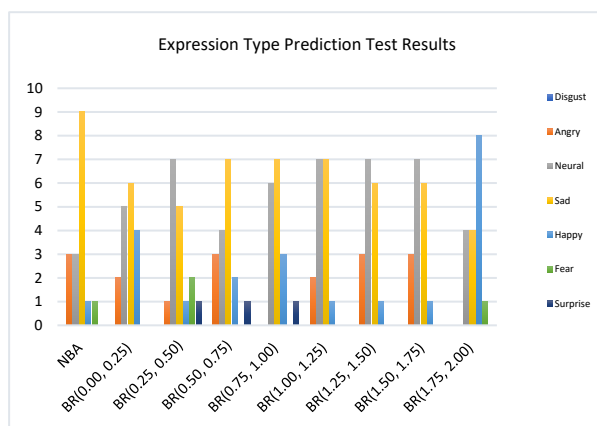


Fig. 17. Expression Type Prediction Test Results MobileNetFLCV.

Fig. 17 shows the results of expression types generated in 17 samples used in the prediction test process with the MobileNet-FLCV architecture based on the techniques used. In addition, based on the figure, it shows that the model that can predict the most types of expressions is the model that applies the brightness range (0.25, 0.50), which in 17 samples predicted angry, neutral sad, happy, scared, and surprised expressions. The graph that shows the prediction test results for each sample can be seen in Fig. 18.

Fig. 18 shows the number of facial expression prediction results for each data sample used in the prediction test with MobileNet-FLCV architecture. Data samples that have four types of expression

predictions are in the sixth (S6), seventh (S7), 15<sup>th</sup> (S15), and 17<sup>th</sup> (S17) data samples. In addition, data samples that predicted only one type of expression were in the third data sample (S3), the eighth data sample (S8), the 12<sup>th</sup> data sample (S12), and the 13<sup>th</sup> data sample (S13).

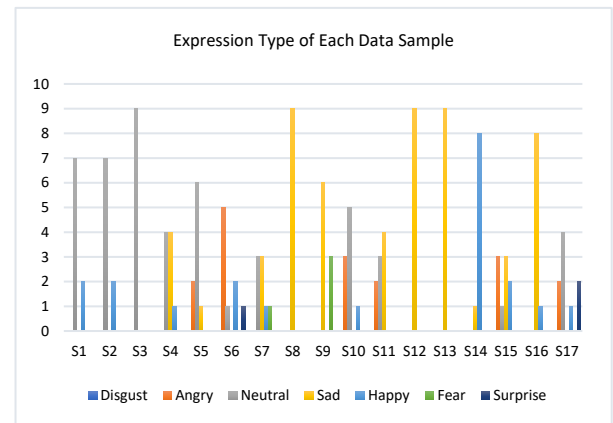


Fig. 18. Expression type of each data sample MobileNet-FLCV.

## 5. Conclusions

This research has successfully applied deep learning techniques using CNN architectures, specifically VGG16 and MobileNet, for masked facial expression classification. This research makes an important contribution to overcoming the challenges of facial expression recognition when wearing masks makes it difficult in social interactions. The results show that the use of VGG16 architecture with cross-validation method (VGG16-FLCV) provides better performance than MobileNet-FLCV architecture in recognizing and classifying masked facial expressions. This shows that in masked facial expression classification applications, VGG16 is superior to MobileNet. The application of data augmentation methods, such as geometric augmentation and brightness augmentation, has helped to improve the performance of CNN models. However, it is important to choose an appropriate brightness range value to obtain optimal results. The experimental results show that on the VGG16-FLCV architecture, the brightness range (1.00, 1.25) provides the best accuracy with a training accuracy of 81.73% and 70.71% validation accuracy.

The application of brightness augmentation on the MobileNet-FLCV architecture does not provide comparable performance to VGG16-FLCV. From the results of this study, it is found that the optimal application of brightness range on VGG16-FLCV architecture is in the darkness category with the brightness ranges (0.25, 0.50), (0.50, 0.75) and (0.75,

1.00) and in the brightness category with brightness range (1.00, 1.25). In addition, this study found that the MobileNet-FLCV architecture with brightness ranges (0.25, 0.50), (0.50, 0.75), (0.75, 1.00), (1.00, 0.25), and (1.25, 1.50) can be used as an alternative brightness range without experiencing a significant decrease in accuracy. Therefore, the results of this study provide a reference for the selection of the right brightness range in the application of data augmentation in CNN models, especially in the context of masked facial expression classification. This information is potentially useful for the development of more effective masked facial expression recognition technology and to support social interactions in pandemics or environments with extensive use of masks.

This research possesses several weaknesses and shortcomings. Therefore, the following suggestions can be used as a reference for further research and development, including applying other types of pre-trained model architectures in implementing transfer learning principles for the classification or prediction of masked facial expression types. Age parameters can be added to the process of predicting the type of masked facial expression to predict the type of masked facial expression and age of masked individuals. The development of the CNN model in predicting masked facial expressions is needed so that the process of detecting masked facial expressions can be done in real-time or be used for developing applications to detect masked facial expressions.

## References

- Agrawal, A. & Mittal, N. (2019). Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy. *The Visual Computer*, 2. <https://doi.org/10.1007/s00371-019-01630-9>
- Berrar, D. (2018). Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, 1–3(April), 542–545. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>
- Castellano, G., De Carolis, B. & Macchiarulo, N. (2021). Automatic emotion recognition from facial expressions when wearing a mask. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3464385.3464730>
- Cheng, X., Lu, J., Member, S. & Yuan, B. (n.d.). Face Segmentor-Enhanced Deep Feature Learning for Face Recognition. 1–14.
- Choe, J. & Shim, H. (n.d.). Attention-based Dropout Layer for Weakly Supervised Object Localization. 2219–2228.
- Cotter, S.F. (n.d.). MobiExpressNet: A Deep Learning Network for Face Expression Recognition on Smart Phones. 2020 IEEE International Conference on Consumer Electronics (ICCE), 1–4.
- Ding, H., Zhou, P. & Chellappa, R. (2020). Occlusion-adaptive deep network for robust facial expression recognition. *IJCB 2020 - IEEE/IAPR International Joint Conference on Biometrics*. <https://doi.org/10.1109/IJCB48548.2020.9304923>
- Farkhod, A. & Abdusalomov, A.B. (2022). Development of Real-Time Landmark-Based Emotion Recognition CNN for Masked Faces.
- Genç, Ç., Colley, A., Löchtefeld, M. & Häkkinä, J. (2020). Face mask design to mitigate facial expression occlusion. *Proceedings - International Symposium on Wearable Computers, ISWC*, 40–44. <https://doi.org/10.1145/3410531.3414303>.
- Grundmann, F., Epstude, K. & Id, S.S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. 1–18. <https://doi.org/10.1371/journal.pone.0249792>.
- Kandel, I., Castelli, M., & Manzoni, L. (2022). Brightness as an Augmentation Technique for Image Classification. *Emerging Science Journal*, 6(4), 881–892. <https://doi.org/10.28991/ESJ-2022-06-04-015>.
- Li, J., Jin, K., Zhou, D., Kubota, N., & Ju, Z. (2020). Attention Mechanism-based CNN for Facial Expression. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2020.06.014>
- Li, Z., Kamnitsas, K., & Glocker, B. (2021). Analyzing Overfitting Under Class Imbalance in Neural Networks for Image Segmentation. 40(3), 1065–1077.
- Merghani, W., & Yap, M. H. (n.d.). Adaptive Mask for Region-based Facial Micro-Expression Recognition. 2–7.
- Pei, Z., Xu, H., Zhang, Y., Guo, M., & Yang, Y. (2019). Face Recognition via Deep Learning Using Data Augmentation Based on Orthogonal Experiments. 1–16. <https://doi.org/10.3390/electronics8101088>.
- Rice, L., Wong, E., & Kolter, J. Z. (2020). Overfitting in adversarially robust deep learning. 37th International Conference on Machine Learning,

ICML 2020, PartF16814, 8049–8074.

Shinde, P.P. & Shah S. (2018). A Review of Machine Learning and Deep Learning Applications. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). IEEE, 1–6.

<https://doi.org/10.1109/ICCUBEA.2018.8697857>.

Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F. & Pinheiro, P.R. (2020). CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection. IEEE Access, 8, 91916–91923.

<https://doi.org/10.1109/ACCESS.2020.2994762>.

Yang, B. (2021). Face Mask Aware Robust Facial Expression Recognition. 240–244.

Yang, B., Wu, J., Ikeda, K., Hattori, G., Sugano, M., Iwasawa, Y., & Matsuo, Y. (2022). Face-mask-aware Facial Expression Recognition based on Face Parsing and Vision Transformer. Pattern Recognition Letters, 164, 173–182.

<https://doi.org/10.1016/j.patrec.2022.11.004>.

Zhang, H., Zhang, L., & Jiang, Y. (n.d.). Overfitting and Underfitting Analysis for Deep Learning Based End-to-end Communication Systems. 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP), 1–6.

## AUTHOR BIOGRAPHIES



### **Desty Mustika Ramadhan**

completed her undergraduate education in informatics at Siliwangi University. During her study period, she actively participated in the independent study program (MSIB) at the independent campus at Bangkit Academy, which took her on the machine learning path.



**Husni Mubarak** completed his master's education at the Bandung Institute of Technology in 2011. He is currently a researcher and lecturer in informatics engineering at Siliwangi University, Tasikmalaya. Research areas include Artificial

Intelligence, Machine Learning, Recognition Systems, Intelligent Systems, and Pattern Recognition.



**Rianto** completed his master's education at the Bandung Institute of Technology in 2015. He is currently a researcher and lecturer in informatics engineering at Siliwangi University, Tasikmalaya. Research areas include Software Engineering,

Artificial Intelligence, Data Mining, Data Processing, Text Mining, Sentiment Analysis, Information Science.

# A systematic meta-analysis on the role of artificial intelligence and machine learning in detection of gynecological disorders

Jyoti I. Nandalwar<sup>1\*</sup>, Pradeep M. Jaeandhiya<sup>2</sup>

<sup>1</sup>CSE Dept., Shree Siddheshwar Women's College of Engineering, Solapur, Maharashtra, India

<sup>2</sup>Department Principal, P.R. Pote Patil College of Engineering and Management, Amravati, Maharashtra, India

\* Corresponding author E-mail: jyotinandalwar@gmail.com

(Received 29 July 2024; Final version received 11 October 2024; Accepted 21 October 2024)

## Abstract

Globally, one of the major concerns in women's health issues is gynecological disorders such as cancer, which needs to be observed at its early stage. With traditional approaches, it is quite difficult to detect such disorders at its early stages. Therefore, more advanced tools need to be integrated. This paper focuses the advancements of artificial intelligence (AI) and machine learning (ML), exploring their potential in the early detection and diagnosis of these disorders. This paper presents a systematic meta-analysis of AI/ML approaches employed in the diagnosis of gynecological disorders using medical imaging modalities such as magnetic resonance imaging (MRI), ultrasound, etc. The flow for systematic meta-analysis is based on designing the research objective, selection and searching approach with inclusion and exclusion strategy; quality assessment is performed then; and finally, discussion of interpretations is also presented. This paper investigates how ML algorithms can extract characteristics from MRI images and how to use ML to extract and recognize the features from medical images such as MRI, ultrasound, computed tomography (CT) scans, etc. for early detection of gynecological tumors and provision of more personalized risk assessment. However, it is observed that there is a significant impact of advancement of AI/ML on medical technology in the future. Therefore, this paper presents a significant contribution for future medical applications and innovations.

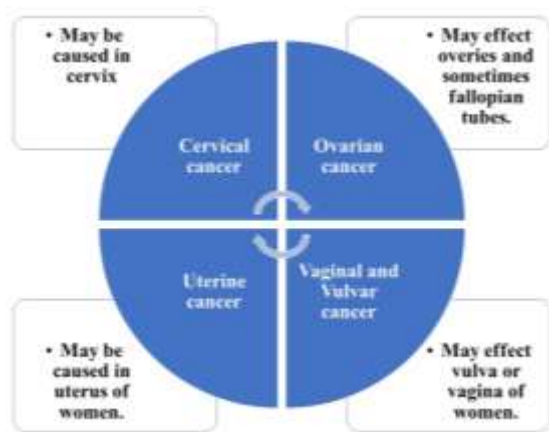
*Keywords:* Artificial Intelligence, Gynecological Cancer, Machine Learning, MRI, Ultrasound

## 1. Introduction

Oncologists in practice face an understanding deficit as a result of the exponential growth of information about cancer together with the fast development of human society (Bhattacharjee et al., 2017; Prapty & Shitu, 2020). Researchers have demonstrated that it is quite challenging for physicians to handle clinical workloads in shorter timeframe to gain professional expertise (Denny et al., 2019; Mehrotra et al., 2011). More individuals may take advantage of societal investments in research and development, physicians can promote and embrace innovative prediction, diagnostic, and treatment procedures based on the best available evidence.

Artificial intelligence (AI) (Constantinou et al., 2009; Liu et al., 2021) has entered the medical field for academic advancements, followed by an immediate introduction of new tools and techniques. One of the most challenging instances for oncology is the identification of gynecological malignancy. The many malignancy forms, every one of which is named after the body part in which it initially manifests, are shown in Fig. 1 (Basij et al., 2018; Chauhan & Singh, 2021). The goal of AI is to develop intelligent machines that can emulate human cognitive processes (Kajala & Jain, 2020; Ray, 2019). By increasing precision, effectiveness, and scalability, these developments have an opportunity to transform the diagnosis of

gynecological cancer (Hou et al., 2022). Over the past decades, AI has been developed to be used for scientific application and medical diagnosis because AI can learn from patterns or knowledge from data and predict accurate outcomes, as compared to conventional techniques (Akazawa & Hashimoto, 2021; Shrestha et al., 2022). Therefore, several types of AI/ML have been used by researchers for early diagnosis of gynecological cancer. For this medical condition, imaging analysis is generally used. Gynecological cancers can be easily detected using radiological imaging techniques such as magnetic resonance imaging (MRI), computed tomography (CT) scans, histopathology, and ultrasounds (Zhou et al., 2020). This approach advances the diagnosis performance in women's health also. Apart from imaging technologies, other parameters such as metabolic testing, polycystic ovarian disease (PCOD), and polycystic ovary syndrome (PCOS) can also be used for diagnosis. These AI/ML algorithms can detect risk factors and suggest patient care. While AI/ML has massive potential for detecting gynecological cancer, there are still several research gaps due to inadequate data quality and availability.



**Fig. 1.** Types of gynecological cancer

The primary goal of the paper is to give a thorough analysis of the most recent research on AI/ML for automated detection of gynecological cancer. Other objectives are as follows:

To investigate and study the advancement of AI tools and techniques in integrating and analyzing diverse datasets, such as medical imaging, health parameters, and clinical information for enhanced gynecological cancer detection.

To investigate the potential of ML-based approaches in early detection and classification of gynecological malignancies.

To contribute to the body of knowledge in the field of gynecological oncology by providing insights into the potential applications and limitations of AI and ML techniques in cancer detection.

## 2. Literature Review

In the early 1950s, AI first became a reality. The invention of intelligent robots with human-like capabilities and responses, or AI, has been universally acknowledged as having a significant impact on the industrial sector (Kasture et al., 2021; Tanaka et al., 2016). Medical professionals increasingly integrated AI into the field of medicine after realizing its significance. Therefore, AI can be broadly used in medical field. There are a lot of recent instances of AI being used in medicine. By generating a new field of study to be exploited for precision medicine, ML and radiomics are revolutionizing radiology and medicine. Gynecologic oncologists are hesitant to fall behind as AI continues to advance in the world of medicine. Medical imaging data from MRI and ultrasound scans may be analyzed by AI algorithms to help with the identification and characterization of gynecological cancers. Radiologists may now conduct more precise evaluations and identify small abnormalities that could be difficult to see visually because of the advancements made in AI approaches, notably deep learning. Researchers have reported that recent advancements of AI/ML had revolutionized the assessment and management of risk of gynecological cancer and provide personalized treatment plans (Hu et al., 2023; Lingappa & Parvathy, 2023).

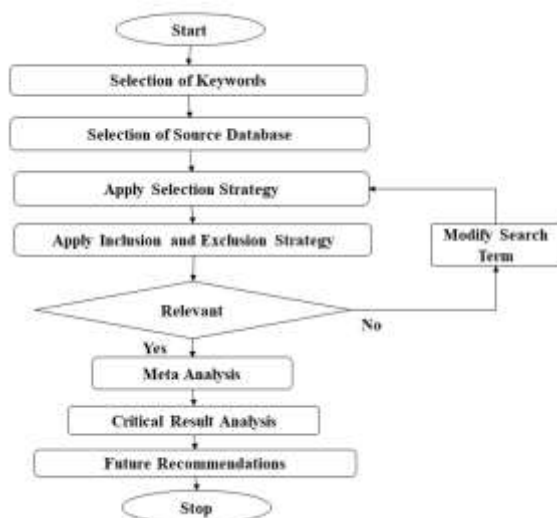
Several studies have highlighted the impact of AI/ML on gynecological cancer diagnosis and risk assessment. For instance, deep learning models, such as Convolutional Neural Networks (CNNs), have shown great promise in the detection of early-stage tumors from MRI and ultrasound scans. Other machine learning techniques, such as Support Vector Machines (SVMs) and Random Forests, have also been employed to improve the classification of cancerous versus non-cancerous tissues. These advancements are transforming the way radiologists and oncologists approach cancer diagnosis, offering more personalized treatment plans tailored to the specific needs of each patient. Deep learning models are highly effective in image analysis but it requires



large datasets of labeled medical images to achieve optimal performance. This poses a challenge in gynecological oncology but it has limitation that there is lack of labeled data. Additionally, widespread clinical adoption is another issue related to the interpretability of ML models. Traditional machine learning models, while easier to interpret, may not achieve the same level of diagnostic accuracy as seen in more complex AI systems like deep learning networks.

### 3. Method Used

In this section, the methodology adopted for performing systematic meta-analysis for the identification of the role of AI/ML in the detection of gynecological tumors or cancer. For this systematic meta-analysis is performed. The flowchart of paper selection for meta-analysis is presented below in Fig. 2.



**Fig. 2.** Methodology used for systematic meta-analysis

The flowchart outlines the systematic meta-analysis process for identifying the role of AI/ML in the detection of gynecological tumors or cancer. The process begins with defining the objective of the systematic meta-analysis. In this paper, the objective is to present a systematic review and meta-analysis for role of AI/ML for automatic detection and correlation of gynecological cancers such as endometrium cancer, cervical cancer, ovarian cancer, etc. This paper aims to analyze the imaging technologies for women health diagnosis by detecting gynecological cancer. To

achieve these objectives, relevant keywords are chosen to perform the literature search. These keywords could include terms such as “AI in gynecology,” “machine learning for cancer detection,” “gynecological tumors,” “medical imaging for gynecological tumors detection,” and “AI for gynecological diagnosis.” After determining the keywords, suitable scientific databases were selected for the literature search. Common databases may include IEEE Xplore, Science direct, Springer, Wiley, etc. that contain relevant studies on AI/ML and gynecological cancer detection. A selection strategy was applied to filter relevant research papers from the selected databases. This may involve screening by title, abstract, and keywords to identify studies that align with the research objectives. Then, inclusion and exclusion strategy was applied to further refine the selection. Inclusion criteria could involve selecting studies focused on AI/ML methods for medical imaging in gynecology, while exclusion criteria might omit studies that lack sufficient data or focus on unrelated medical conditions. The filtered studies were assessed for relevance to the research topic. If a study is deemed not relevant, the search is modified to refine the keywords or selection criteria. If the study is relevant, it is included for further analysis. Relevant studies were subjected to meta-analysis and analysis to draw conclusions about the overall impact of AI/ML techniques on the diagnosis of gynecological disorders. After the meta-analysis, the results were critically analyzed to identify the strengths and limitations of the AI/ML approaches.

### 4. Detection Using Machine Learning Based on MRI Images

For medical diagnosis, MRI findings are considered one of the most important inputs. To better understand the patterns in MRIs, ML extracts characteristics more precisely (Subramanian et al., 2023). These characteristics may be the tissue texture, shape, intensity, and spatial connections. These need to be identified and learned properly to distinguish between healthy and malignant regions in MRI images (Baydoun et al., 2021). To provide the capability of early prediction, ML models may be trained to identify patterns and distinguish certain biomarkers (Zhang & Han, 2020). Additionally, segmentation methods can accurately identify the tumor borders by applying ML techniques that will provide valuable information to doctors for planning, tracking, and treating

gynecological disorders (Guo et al., 2019; Visalaxi et al., 2021). Soğukkuyu and Ata (2022) proposed an ensemble strategy for predicting the risk of cervical cancer. Keymasi et al. (2018) categorized pap-smear pictures using various ML techniques to increase prediction accuracy. KNN, SVM, and multilayer perceptron (MLP) are three ML approaches that are combined in the ensemble methodology. Jiang et al. (2021) used multi-parametric MRI data to construct radiomics algorithms that are based on deep learning, which are used for the detection of cervical cancer. Wang et al. (2021) developed a methodology based on deep learning that distinguishes between malignant and benign ovary lesions using CNN as opposed to using conventional MR imaging. A single institution divided 451 patients' 545 lesions—379 normal and 166 malignant—into 7:2:1 training, validation, and evaluation sets. Ghoneim et al. (2020) presented a technique for the identification and classification of cervical cancer cells that is based on CNN. Following this step, the input images are assigned categories using a classifier that is driven by an extreme machine learning (ELM). Transfer learning and fine-tuning are the two methods that are employed to use the CNN architecture. In addition to the ELM, classifiers based on autoencoders and multilayer perceptrons are also being researched as potential replacements. Ratul et al. (2022) carried out a logical analysis to show the efficiency of the MLP method with default hyperparameters and obtained 93.33% prediction accuracy. Wang et al. (2023) offered a unique optical biopsy technique to help surgeons quickly and reliably diagnose ovarian cancer. Khuriwal and Mishra (2018)

showed how AI may be used with the UCI Database to identify breast cancer. Kurnianingsih et al. (2019) used a mask regional CNN (Mask R-CNN) on the pap smear dataset to assess the cancerous cells. Arora et al. (2021) used SVM and achieved an accuracy of 95%. The author also used the Gaussian filter for image denoising. Bnoui et al. (2021) suggested an ensemble preprocessing technique to boost a CNN's classification accuracy for cervical cancer. Table 1 presents the overview of recent research on MRI images for detection of gynecological cancer.

### 5. Detection Using Machine Learning Based on Ultrasound Images

In gynecology, ultrasound imaging is a frequently used diagnostic technique for the identification and assessment of gynecological malignancies. To better analyze ultrasound data and diagnose gynecological cancer, ML algorithms have shown potential (Zhang & Han, 2020). Algorithms are used in ML-based techniques to categorize anomalies, extract useful information from ultrasound pictures, and assist in the early diagnosis of gynecological cancers (Zhang et al., 2023). ML algorithms can examine ultrasound images by identifying characteristics like vascularity, echogenicity, texture, and spatial connections (Ruchitha et al., 2021). Studies have revealed that ML-based methods may improve diagnostic precision by using ultrasound images (Marques et al., 2019). Ultrasound images can be used in environments with minimal resources due to their portability and accessibility (Behboodi et al., 2021). Arezzo et al. (2022) used ML to detect cancer on

**Table 1.** Recent literature on MRI-based gynecological cancer detection integrated with AI/ ML

Ref.	Year	Methods used	Imaging modality	Type of cancer	Result
Soğukkuyu & Ata (2022)	2022	Multiple ML methods	MRI	Cervical cancer	Accuracy = 97%
Keymasi et al. (2018)	2018	KNN, SVM	MRI	Cervical cancer	Accuracy = 97.83%
Jiang et al. (2021)	2021	Deep learning-based radiomic methods	MRI	Early-stage cervical cancer	AUC = 91.1% Sensitivity = 88.1%
Wang et al. (2021)	2021	CNN	MRI	Ovarian cancer	Accuracy = 87% Sensitivity = 75%
Ghoneim et al. (2020)	2020	CNN, MLP	MRI	Cervical cancer	Accuracy = 99.5%
Ratul et al. (2022)	2022	KNN, DTC, SVM, RFC, MLP	MRI	Cervical cancer	MLP performed best with accuracy = 93.33%
Wang et al. (2023)	2023	End-to-end deep learning	MRI	Ovarian cancer	Accuracy = 99.7%
Arora et al. (2021)	2021	SVM	MRI	Cervical cancer	Accuracy = 95%

ultrasound data. Gao et al. (2022) intended to create a deep CNN system that automates ultrasound image interpretation and makes ovarian cancer detection easier than with current techniques. Srivastava et al. (2020) obtained sample ultrasound pictures of the ovaries from various women and identified the presence or absence of ovarian cysts. The standard VGG-16 model is used in the proposed research and is tweaked using an exclusive dataset of ultrasound images. A 16-layer DL-NN trained on the ImageNet dataset is a VGG-16 model. The last four layers of the VGG-16 network are changed to adjust the network. Kiruthika et al. (2020) presented an artificial neural network to construct an intelligent automated detection and ovarian categorization for ovary detection was given three texture characteristics. Zhou et al. (2021) evaluated the usefulness of tumor feature extraction on DBN for cancer of the cervical cavity patient diagnosis and to accomplish a smart assessment of the impacts of therapy and cervical detection of cancer. This technique was then used to analyze tumors automatically using a DBN architecture.

The proposed framework for tumor extraction of features based on the DBN was shown to have a superior accuracy of 86.36%, sensitivity of 83.33%, and specificity of 87.50 %. Taleb et al. (2022) have shown the capability of ML for accurate recognition of

ovarian cancer and its stages. The majority of current studies on ovarian cancer employ a single categorization model, which has poor diagnostic efficacy. Chen et al. (2021) proposed a 3D CNN based on a domain-knowledge-guided temporal attention module and a channel attention module. On a dataset of 221 breast-CEUS patients, the author validated the model. Hyun et al. (2020) proposed a 4-layer CNN to identify MB signatures without causing any damage. Goudarzi et al. (2023) studied the segmentation approach on ultrasound images with a Dice Score Coefficient (DSC) of 0.940.08 and 0.920.06, respectively. The CutMix augmentation technique enhances the generalization performance of the proposed CNN, which is tuned for accurate automated segmentation of tissue layers. Table 2 presents the overview of recent research on ultrasound images for detection of gynecological cancer.

## 6. Detection Using Machine Learning Based on Metabolic Parameters

Machine learning algorithms may use a variety of health factors in addition to medical imaging to diagnose gynecological cancer. These health indicators include blood test results, metabolic test results, PCOS, PCOD, and other pertinent clinical data (Coffin et al., 2023). To identify risk variables,

**Table 2.** Recent literature on ultrasound-based gynecological cancer detection in integration with AI/ML

Ref.	Year	Method	Imaging modality	Type of cancer	Result
Arezzo et al. (2022)	2022	RF, LR, KNN	Ultrasound	Ovary cancer	Accuracy = 93.7% Precision = 90% Recall = 90%
Gao et al. (2022)	2022	D-CNN	Ultrasound	Ovary cancer	AUC = 91.1% Accuracy = 86.9%
Srivastava et al., (2020)	2020	VGG-16	Ultrasound	Ovary cancer	Accuracy = 92.11%
Kiruthika et al. (2020)	2020	ANN	Ultrasound	Ovary cancer	Accuracy of 96%
Zhou et al. (2021)	2021	DBN	Ultrasound	Cervical cancer	Accuracy = 86.36% Sensitivity = 83.3% Specificity = 87.50%
Taleb et al. (2022)	2022	SVM, KNN	Ultrasound	Ovary cancer	Accuracy = 98.1% and 97.16%
Chen et al. (2021)	2021	3D CNN	Ultrasound	Breast cancer	Sensitivity of 97.2% and an accuracy of 86.3%.
Hyun et al. (2020)	2020	4-Layer CNN	Ultrasound	Breast cancer	Generalized contrast-to-noise ratio (GCNR) of 0.93 and Kolmogorov-Smirnov statistic (KSS) of 0.86
Goudarzi et al. (2023)	2023	Gated Shape CNN	Ultrasound	Breast cancer	DSC = 94% and 92 %

forecast disease development, and give a tailored risk assessment for gynecological cancer, ML algorithms may combine and evaluate this information. ML algorithms may be used to create prediction models for gynecological cancer risk assessment using health factors like blood test results and metabolic test results. ML models may establish relations among certain biomarkers, patterns, or characteristics that can increase the risk of developing gynecological malignancies (Harish et al., 2023; Tiwari et al., 2022). ML algorithms are capable of enhancing the identification and management of gynecological cancer by analyzing the health metrics or genetic matrices to present personalized risk assessments and treatment recommendations (Harish et al., 2023). ML can analyze medical data to identify risk factors and early signs of cancer and also enable medical experts to provide early interventions like lifestyle adjustments or preventive measures. Bharati et al. (2020) discussed the data-driven approach to diagnosis of PCOS in women. The model has achieved 91.01% of accuracy and 90% of recall value. Poorani and Khilar (2023) categorized whether a woman has PCOS. Denny et al. (2019) provided a method for the timely diagnosis and prognosis of PCOS by making use of clinical and nutritional markers that are ideal and minimal but still helpful in predicting the presence of the condition. The 541 women who participated in this study provided the data sets that were necessary for the construction of this framework. A variety of different ML techniques are used to categorize PCOS utilizing the accumulated set of characteristics that were altered using principal component analysis (PCA). Chitra et al. (2023) diagnosed PCOS via ultrasound images by using transfer learning approaches such as Alexnet and Inception. Harish et al. (2023) identified that PCOS is a serious condition that affects females when their ovaries are fertile, between the ages of 15 and 45. This disease affects 5–10% of reproductive-age females. Despite the difficulty in fully treating this condition, PCOS-affected women may minimize their symptoms by getting the right amount of exercise, eating well, and maintaining a healthy BMI. Random oversampling triumphs when comparing the two equally weighted approaches using accuracy. An enhanced AI classifier for PCOS diagnosis was developed using 594 ovarian ultrasonography (USG) images by (Suha & Islam, 2022). The proposed technique outperforms previous ML-based methods in accuracy and training time. Utilizing the recommended expanded method, the

"VGGNet16" pre-trained algorithm uses a CNN architecture as an extractor of features and a stacked ensemble algorithm with the "XGBoost" meta-learner as an image classification. Khanna et al. (2023) have shown an AI method for predicting PCOS in patients with fertile patients utilizing heterogeneous ML/DL approaches. The author investigated a 541-patient open-source dataset. Swapnarekha et al. (2023) proposed a predicting model based on random oversampling that has proved successful in resolving the issue of class imbalance. In this method, the optimal model hyperparameters are selected via Bayesian optimization. Wang et al. (2022) proposed a decision tree-based PCOS-linked cancer detection algorithm. Based on the obtained SMFs, a decision tree is built, and practical simulations are run on separate internal and external cohorts. Bharati et al. (2022) presented the statistical approach for women's PCOS diagnosis. When the traits are graded, the ratio of luteinizing hormone (LH) to follicle-stimulating hormone (FSH) is shown to be the most important factor. While features are being chosen and eliminated, the cross-validation approach is used. Among the classifiers utilized on the dataset are voting hard, voting soft, and CatBoost. Prasher and Nelson (2023) identified the hormone that is most often seen globally is PCOS. The ovaries generate a large number of microscopic fluid-filled sacs known as follicles, which are the root cause of PCOS. The ovaries did not always release the eggs, as was to be anticipated. One of the best ways to diagnose PCOS early and develop a treatment plan for people with this condition is to look for numerous follicles on USG scans. Nasim et al. (2022) predicted the PCOS by applying ML approaches. Based on the CS-PCOS mechanism, a unique feature selection method is suggested. Prolactin, blood pressure (systolic and diastolic), and pregnancy are the key indicators with significant influence on PCOS prediction. Through the early discovery of PCOS, the work assists the medical community in reducing the miscarriage rate and offering women a remedy. Table 3 presents the overview of recent research on metabolic parameters for detection of gynecological cancer.

## 7. Discussion

The current study provides insights into the role of AI/ML techniques in the detection of gynecological cancers through MRI, ultrasound, and metabolic parameters. The studies used for meta-analysis are

**Table 3.** Recent literature on metabolic parameters-based gynecological cancer detection in integration with AI/ ML.

Ref.	Year	Methods	Metabolic test	Type of cancer	Result
Bharati et al. (2020)	2020	LR	PCOS	Ovary cancer	Accuracy = 91.01%
Denny et al. (2019)	2019	KNN, SVM	PCOS	Ovary cancer	Accuracy = 89.02%
Chitra et al. (2023)	2023	Inception V3, Resnet50, VGG16 and Hybrid Models	PCOS	Ovary cancer	Accuracy = 93%
Harish et al. (2023)	2023	SVM, XGBOOST, LR, KNN, RF	Polycystic ovary syndrome	Ovary cancer	Accuracy = 96%
Swapnarekha et al. (2023)	2023	SVM, Genetic algorithm, MLP, ELM	Polycystic ovary syndrome	Ovary cancer	Accuracy = 99.31%
Wang et al. (2022)	2022	Decision Tree	PCOS	Ovary cancer	AUC = 96.7%
Bharati et al. (2022)	2022	Ensemble Learning	Polycystic ovary syndrome	Ovary cancer	Accuracy = 91.12%
Prasher & Nelson (2023)	2023	D-CNN	Polycystic ovary syndrome	Ovary cancer	Accuracy = 99.4%
Nasim et al. (2022)	2022	Generative NB	Polycystic ovary syndrome	Ovary cancer	Accuracy = 100%

selected to effectively demonstrate the role of ML and deep learning models for cancer detection.

For instance, high accuracy has been achieved by CNN-based models, as seen in studies by Wang et al. (2021) and Ghoneim et al. (2020). This shows the trend of using CNN for image-based diagnosis. The use of ensemble methods, such as those proposed by Soğukkuyu and Ata (2022) and Keymasi et al. (2018) show promising improvement in prediction accuracy. This indicates a trend toward using hybrid models to address the limitations of single-model approaches. Our review of the studies using MRI-based detection techniques shows that deep learning models like CNNs outperform consistently over traditional ML approaches such as KNN, SVM, and MLP in detecting and classifying gynecological cancers. This is evident in studies by Jiang et al. (2021) and Wang et al. (2021) where CNN models demonstrated its better performance compared to non-deep learning methods. Similarly, ultrasound-based detection techniques also shows the benefit of usage of deep learning. Both Gao et al. (2022) and Srivastava et al. (2020) showed that CNN models can improve the accuracy of ovarian cancer detection compared to traditional methods. However, Taleb et al. (2022) and Chen et al. (2021) used the attention modules or domain-guided approaches that can further enhance performance. In terms of metabolic parameter-based detection, researchers showed that integration of metabolic data

with ML techniques provided a significant improvement in cancer risk assessment. Bharati et al. (2020) demonstrated how ML algorithms could effectively combine metabolic data to predict PCOS-related ovarian cancer with high accuracy. This trend highlights the growing importance of using multi-modal data to provide a more comprehensive view of cancer risk and progression.

However, several limitations were also identified:

Many of the studies rely on relatively small datasets. MRI and ultrasound datasets are often limited in size. To mitigate these limitations, there is a need to develop larger and diverse datasets that can help to improve the generalizability of AI/ML models.

While deep learning models such as CNN are “black-box” in nature that presents a challenge for clinical adoption, Ghoneim et al. (2020) and Chen et al. (2021) incorporated more interpretable models such as attention models. Hybrid models that combine multiple AI/ML approaches could yield better outcomes.

## 8. Practical Implication of Findings

The meta-analysis presented in this paper shows the significant practical implications for the integration of AI/ML techniques in clinical applications. AI/ML models such as deep learning algorithms like CNNs have shown good accuracy for identifying gynecological tumors from MRI and

ultrasound images. These models can serve as decision support tools for radiologists and gynecologists for early detection of cancer. By integrating metabolic parameters and health indicators, AI/ML models can provide more personalized risk assessments and treatment recommendations. The use of AI with imaging techniques is particularly promising for resource-limited environment. Such applications can be used in remote areas to facilitate early detection and diagnosis of gynecological cancers. This will reduce the need for immediate specialist intervention. AI/ML systems can minimize diagnostic errors caused by human fatigue, inexperience, or oversight. However, it has some potential barriers also such as lack of large, diverse, and high-quality datasets, hybrid models are black-box in nature, resource limitations, complexity, etc.

## 9. Conclusion

The use of AI/ML approaches to improve the detection and treatment of gynecological malignancies has gained increasing attention in the past few years. Therefore, the goal of this paper is to describe the current level of AI research for the diagnosis of gynecological tumors using imaging technologies and health metabolic parameters. This paper presents a systematic meta-analysis that explored these aspects. The paper identified that deep learning model as well as hybrid or ensemble learning models outperforms better diagnosis either it is MRI imaging or ultrasound imaging approach. Whereas, on cancer risk assessment, metabolic parameters were also considered and its diagnosis hybrid model also outperforms well. Whereas potential barriers were also identified in this research, such as lack of large, diverse, and high-quality datasets, hybrid models are black-box in nature, resource limitations, complexity, etc. These aspects will need to be focused in future for improvement.

## References

- Adweb, K. M. A., Cavus, N., & Sekeroglu, B. (2021). Cervical Cancer Diagnosis Using Very Deep Networks Over Different Activation Functions. *IEEE Access*, 9, 46612–46625. <https://doi.org/10.1109/ACCESS.2021.3067195>
- Akazawa, M., & Hashimoto, K. (2021). Artificial intelligence in gynecologic cancers: Current status and future challenges - A systematic review. *Artificial Intelligence in Medicine*, 120, 102164. <https://doi.org/10.1016/j.artmed.2021.102164>
- Arezzo, F., Cormio, G., La Forgia, D., Santarsiero, C. M., Mongelli, M., Lombardi, C., Cazzato, G., Cincinelli, E., & Loizzi, V. (2022). A machine learning approach applied to gynecological ultrasound to predict progression-free survival in ovarian cancer patients. *Archives of Gynecology and Obstetrics*, 306(6), 2143–2154. <https://doi.org/10.1007/s00404-022-06578-1>
- Arora, A., Tripathi, A., & Bhan, A. (2021). Classification of Cervical Cancer Detection using Ma-chine Learning Algorithms. 2021 6th International Conference on Inventive Computation Technologies (ICICT), 827–835. <https://doi.org/10.1109/ICICT50816.2021.9358570>
- Basij, M., Yan, Y., Alshahrani, S., Winer, I., Burmeister, J., Dominello, M., & Mehrmohammadi, M. (2018). Development of an Ultrasound and Photoacoustic Endoscopy System for Imaging of Gynecological Disorders. 2018 IEEE International Ultrasonics Symposium (IUS), 1–4. <https://doi.org/10.1109/ULTSYM.2018.8579788>
- Baydoun, A., Xu, K. E., Heo, J. U., Yang, H., Zhou, F., Bethell, L. A., Fredman, E. T., Ellis, R. J., Podder, T. K., Traughber, M. S., Paspulati, R. M., Qian, P., Traughber, B. J., & Muzic, R. F. (2021). Synthetic CT Generation of the Pelvis in Patients with Cervical Cancer: A Single Input Approach Using Generative Adversarial Network. *IEEE Access: Practical Innovations, Open Solutions*, 9, 17208–17221. <https://doi.org/10.1109/access.2021.3049781>
- Behboodi, B., Rivaz, H., Lalondrelle, S., & Harris, E. (2021). Automatic 3D Ultrasound Segmentation of Uterus Using Deep Learning. *IEEE International Ultrasonics Symposium, IUS*, 3–6. <https://doi.org/10.1109/IUS52206.2021.9593671>
- Bharati, S., Podder, P., & Hossain Mondal, M. R. (2020). Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Algorithms. 2020 IEEE Region 10 Symposium (TENSYP), 1486–1489. <https://doi.org/10.1109/TENSYP50017.2020.9230932>
- Bharati, S., Podder, P., Mondal, M. R., Prasath, S., & Gandhi, N. (2022). Ensemble Learning for Data-Driven Diagnosis of Polycystic Ovary Syndrome. 1250–1259. [https://doi.org/10.1007/978-3-030-96308-8\\_116](https://doi.org/10.1007/978-3-030-96308-8_116)



- Bhattacharjee, S., Singh, Y. J., & Ray, D. (2017). Comparative performance analysis of machine learning classifiers on ovarian cancer dataset. 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), 213–218. <https://doi.org/10.1109/ICRCICN.2017.8234509>
- Bhavani, C. H., & Govardhan, A. (2023). Cervical cancer prediction using stacked ensemble algorithm with SMOTE and RFERF. Materials Today: Proceedings, 80, 3451–3457. <https://doi.org/https://doi.org/10.1016/j.matpr.2021.07.269>
- Bnoui, N., Amor, H. Ben, Rekik, I., Rhim, M. S., Solaiman, B., & Amara, N. E. Ben. (2021). Boosting CNN Learning by Ensemble Image Pre-processing Methods for Cervical Cancer Segmentation. 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), 264–269. <https://doi.org/10.1109/SSD52085.2021.9429422>
- Chang, C.-C., Chen, H.-H., Chang, Y.-C., Yang, M.-Y., Lo, C.-M., Ko, W.-C., Lee, Y.-F., Liu, K.-L., & Chang, R.-F. (2017). Computer-aided diagnosis of liver tumors on computed tomography images. Computer Methods and Programs in Biomedicine, 145, 45–51. <https://doi.org/https://doi.org/10.1016/j.cmpb.2017.04.008>
- Chauhan, N. K., & Singh, K. (2021). Impact of Variation in Number of Channels in CNN Classification model for Cervical Cancer Detection. 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 1–6. <https://doi.org/10.1109/ICRITO51393.2021.9596366>
- Chen, C., Wang, Y., Niu, J., Liu, X., Li, Q., & Gong, X. (2021). Domain Knowledge Powered Deep Learning for Breast Cancer Diagnosis Based on Contrast-Enhanced Ultrasound Videos. IEEE Transactions on Medical Imaging, 40(9), 2439–2451. <https://doi.org/10.1109/TMI.2021.3078370>
- Chitra, P., Srilatha, K., Sumathi, M., Jayasudha, F. V., Bernatin, T., & Jagadeesh, M. (2023). Classification of Ultrasound PCOS Image using Deep Learning based Hybrid Models. 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), 1389–1394. <https://doi.org/10.1109/ICEARS56392.2023.10085400>
- Coffin, T., Wray, J., Sah, R., Maj, M., Nath, R., Nauhria, S., Maity, S., & Nauhria, S. (2023). A Review and Meta-Analysis of the Prevalence and Health Impact of Polycystic Ovary Syndrome Among Medical and Dental Students. Cureus, 15(6), e40141. <https://doi.org/10.7759/cureus.40141>
- Constantinou, I. P., Koumourou, C. A., Neofytou, M. S., Tanos, V., Pattichis, C. S., & Kyriakou, E. C. (2009). An integrated CAD system facilitating the endometrial cancer diagnosis. 2009 9th International Conference on Information Technology and Applications in Biomedicine, 1–5. <https://doi.org/10.1109/ITAB.2009.5394424>
- Coşar Soğukkuyu, D. Y., & Ata, O. (2022). Diagnosing Cervical Cancer Using Machine Learning Methods. 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 1–3. <https://doi.org/10.1109/HORA55278.2022.9800033>
- de Carvalho Filho, A. O., Silva, A. C., Cardoso de Paiva, A., Nunes, R. A., & Gattass, M. (2017). Computer-Aided Diagnosis of Lung Nodules in Computed Tomography by Using Phylogenetic Diversity, Genetic Algorithm, and SVM. Journal of Digital Imaging, 30(6), 812–822. <https://doi.org/10.1007/s10278-017-9973-6>
- Denny, A., Raj, A., Ashok, A., Ram, C. M., & George, R. (2019). i-HOPE: Detection and Prediction System For Polycystic Ovary Syndrome (PCOS) Using Machine Learning Techniques. TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), 673–678. <https://doi.org/10.1109/TENCON.2019.8929674>
- Gao, Y., Zeng, S., Xu, X., Li, H., Yao, S., Song, K., Li, X., Chen, L., Tang, J., Xing, H., & Yu, Z. (2022). Articles Deep learning-enabled pelvic ultrasound images for accurate diagnosis of ovarian cancer in China: a retrospective, multicentre, diagnostic study. 179–187. [https://doi.org/10.1016/S2589-7500\(21\)00278-8](https://doi.org/10.1016/S2589-7500(21)00278-8)
- Ghoneim, A., Muhammad, G., & Hossain, M. S. (2020). Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Generation Computer Systems, 102, 643–649. <https://doi.org/https://doi.org/10.1016/j.future.2019.09.015>
- Goudarzi, S., Whyte, J., Boily, M., Towers, A., Kilgour, R. D., & Rivaz, H. (2023). Segmentation

- of Arm Ultrasound Images in Breast Cancer-Related Lymphedema: A Database and Deep Learning Algorithm. *IEEE Transactions on Bio-Medical Engineering*, 70(9), 2552–2563.  
<https://doi.org/10.1109/TBME.2023.3253646>
- Guo, Z., Li, X., Huang, H., Guo, N., & Li, Q. (2019). Deep Learning-Based Image Segmentation on Multimodal Medical Imaging. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 3(2), 162–169.  
<https://doi.org/10.1109/TRPMS.2018.2890359>
- Harish, K. P., Dhivyanchali, M. N., Devi, K. N., Krishnamoorthy, N., Dhana Sree, R., & Dharanidharan, R. (2023). Smart Diagnostic System For Early Detection And Prediction Of Polycystic Ovary Syndrome. 2023 International Conference on Computer Communication and Informatics (ICCCI), 1–6.  
<https://doi.org/10.1109/ICCCI56745.2023.10128560>
- Hou, X., Shen, G., Zhou, L., Li, Y., Wang, T., & Ma, X. (2022). Artificial Intelligence in Cervical Cancer Screening and Diagnosis. *Frontiers in Oncology*, 12, 851367.  
<https://doi.org/10.3389/fonc.2022.851367>
- Hu, D., Jian, J., Li, Y., & Gao, X. (2023). Deep learning-based segmentation of epithelial ovarian cancer on T2-weighted magnetic resonance images. *Quantitative Imaging in Medicine and Surgery*, 13(3), 1464–1477.  
<https://doi.org/10.21037/qims-22-494>
- Hyun, D., Abou-Elkacem, L., Bam, R., Brickson, L. L., Herickhoff, C. D., & Dahl, J. J. (2020). Nondestructive Detection of Targeted Microbubbles Using Dual-Mode Data and Deep Learning for Real-Time Ultrasound Molecular Imaging. *IEEE Transactions on Medical Imaging*, 39(10), 3079–3088.  
<https://doi.org/10.1109/TMI.2020.2986762>
- Jiang, X., Li, J., Kan, Y., Yu, T., Chang, S., Sha, X., Zheng, H., Luo, Y., & Wang, S. (2021). MRI Based Radiomics Approach With Deep Learning for Prediction of Vessel Invasion in Early-Stage Cervical Cancer. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 18(3), 995–1002.  
<https://doi.org/10.1109/TCBB.2019.2963867>
- Kajala, A., & Jain, V. K. (2020). Diagnosis of Breast Cancer using Machine Learning Algorithms-A Review. 2020 International Conference on Emerging Trends in Communication, Control and Computing (ICONC3), 1–5.  
<https://doi.org/10.1109/ICONC345789.2020.9117320>
- Kapruwan, A., Sharma, S., & Goyal, H. R. (2023). Artificial Intelligence Enabled Diagnostic Digital Cytopathology System for Cervical Intraepithelial Neoplasia Detection: Advantages and Challenges. 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), 669–674.  
<https://doi.org/10.1109/ICAIS56108.2023.10073850>
- Kasture, K. R., Sayankar, B. B., & Matte, P. N. (2021). Multi-class Classification of Ovarian Cancer from Histopathological Images using Deep Learning - VGG-16. 2021 2nd Global Conference for Advancement in Technology (GCAT), 1–6.  
<https://doi.org/10.1109/GCAT52182.2021.9587760>
- Keymasi, M., Mishra, V., Aslan, S., & Asem, M. (2018). Theoretical Assessment of Cervical Cancer Using Machine Learning Methods Based on Pap-Smear Test. 1367–1373.  
<https://doi.org/10.1109/IEMCON.2018.8615041>
- Khanna, V. V., Chadaga, K., Sampathila, N., Prabhu, S., Bhandage, V., & Hegde, G. K. (2023). A Distinctive Explainable Machine Learning Framework for Detection of Polycystic Ovary Syndrome. *Applied System Innovation*, 6(2).  
<https://doi.org/10.3390/asi6020032>
- Khuriwal, N., & Mishra, N. (2018). Breast Cancer Diagnosis Using Deep Learning Algorithm. 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 98–103.  
<https://doi.org/10.1109/ICACCCN.2018.8748777>
- Kiruthika, V., Sathiya, S., & Ramya, M.M. (2020). Machine learning based ovarian detection in ultrasound images. *International Journal of Advanced Mechatronic Systems*, 8, 75.  
<https://doi.org/10.1504/IJAMECHS.2020.111306>
- Kurnianingsih, Allehaibi, K. H. S., Nugroho, L. E., Widyawan, Lazuardi, L., Prabuono, A. S., & Mantoro, T. (2019). Segmentation and Classification of Cervical Cells Using Deep Learning. *IEEE Access*, 7, 116925–116941.  
<https://doi.org/10.1109/ACCESS.2019.2936017>
- Li, Y., Chen, J., Xue, P., Tang, C., Chang, J., Chu, C., Ma, K., Li, Q., Zheng, Y., & Qiao, Y. (2020). Computer-Aided Cervical Cancer Diagnosis Using Time-Lapsed Colposcopic Images. *IEEE*

- Transactions on Medical Imaging, 39(11), 3403–3415. <https://doi.org/10.1109/TMI.2020.2994778>
- Lingappa, E., & Parvathy, L. R. (2023). Image Classification with Deep Learning Methods for Detecting and Staging Bone Cancer from MRI. 2023 International Conference on Networking and Communications (ICNWC), 1–8. <https://doi.org/10.1109/ICNWC57852.2023.10127368>
- Liu, X., Xiao, Z., Song, Y., Zhang, R., Li, X., & Du, Z. (2021). A Machine Learning-Aided Framework to Predict Outcomes of Anti-PD-1 Therapy for Patients With Gynecological Cancer on Incomplete Post-Marketing Surveillance Dataset. IEEE Access, 9, 120464–120480. <https://doi.org/10.1109/ACCESS.2021.3107498>
- Marques, S., Carvalho, C., Peixoto, C., Pignatelli, D., Beires, J., Silva, J., & Campilho, A. (2019). Segmentation of gynaecological ultrasound images using different U-Net based approaches. 2019 IEEE International Ultrasonics Symposium (IUS), 1485–1488. <https://doi.org/10.1109/ULTSYM.2019.8925948>
- Mehrotra, P., Chatterjee, J., Chakraborty, C., Ghoshdastidar, B., & Ghoshdastidar, S. (2011). Automated screening of Polycystic Ovary Syndrome using machine learning techniques. 2011 Annual IEEE India Conference, 1–5. <https://doi.org/10.1109/INDCON.2011.6139331>
- Nasim, S., Almutairi, M. S., Munir, K., Raza, A., & Younas, F. (2022). A Novel Approach for Polycystic Ovary Syndrome Prediction Using Machine Learning in Bioinformatics. IEEE Access, 10, 97610–97624. <https://doi.org/10.1109/ACCESS.2022.3205587>
- Nishio, M., & Nagashima, C. (2017). Computer-aided Diagnosis for Lung Cancer: Usefulness of Nodule Heterogeneity. Academic Radiology, 24(3), 328–336. <https://doi.org/https://doi.org/10.1016/j.acra.2016.11.007>
- Nithya, B., & Ilango, V. (2020). Machine Learning Aided Fused Feature Selection based Classification Framework for Diagnosing Cervical Cancer. 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 61–66. <https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00011>
- Poorani, B. & Khilar, R. (2023). Classification of PCOS Using Machine Learning Algorithms Based on Ultrasound Images of Ovaries. 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), 1–7. <https://api.semanticscholar.org/CorpusID:259101230>
- Prapty, A. S., & Shitu, T. T. (2020). An Efficient Decision Tree Establishment and Performance Analysis with Different Machine Learning Approaches on Polycystic Ovary Syndrome. 2020 23rd International Conference on Computer and Information Technology (ICCIT), 1–5. <https://doi.org/10.1109/ICCIT51783.2020.9392666>
- Prasher, S., & Nelson, L. (2023). Follicle Prediction for Polycystic Ovary Syndrome Diagnosis from Ovarian Ultrasound Images using CNN. 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), 789–793.
- Rahman, M. A., Muniyandi, R. C., Islam, K. T., & Rahman, M. M. (2019). Ovarian Cancer Classification Accuracy Analysis Using 15-Neuron Artificial Neural Networks Model. 2019 IEEE Student Conference on Research and Development (SCORED), 33–38. <https://doi.org/10.1109/SCORED.2019.8896332>
- Ratul, I. J., Al-Monsur, A., Tabassum, B., Ar-Rafi, A. M., Nishat, M. M., & Faisal, F. (2022). Early risk prediction of cervical cancer: A machine learning approach. 2022 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 1–4. <https://doi.org/10.1109/ECTI-CON54298.2022.9795429>
- Ray, S. (2019). A Quick Review of Machine Learning Algorithms. 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 35–39. <https://doi.org/10.1109/COMITCon.2019.8862451>
- Ruchitha, P. J., Sai Richitha, Y., Kodipalli, A., & Martis, R. J. (2021). Segmentation of Ovarian Cancer using Active Contour and Random Walker Algorithm. 2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECOT), 238–241. <https://doi.org/10.1109/ICEECOT52851.2021.9707939>
- Senthilkumar, G., Ramakrishnan, J., Frnda, J.,

- Manikandan, R., Gupta, D., Tiwari, P., Mohammad, S., & Mohammed, M. (2021). Incorporating Artificial Fish Swarm in Ensemble Classification Framework for Recurrence Prediction of Cervical Cancer. *IEEE Access*, PP, 1.  
<https://doi.org/10.1109/ACCESS.2021.3087022>
- Shrestha, P., Poudyal, B., Yadollahi, S., E Wright, D., V Gregory, A., D Warner, J., Korfiatis, P., C Green, I., L Rassier, S., Mariani, A., Kim, B., Laugh-lin-Tommaso, S. K., & L Kline, T. (2022). A systematic review on the use of artificial intelligence in gynecologic imaging - Background, state of the art, and future directions. *Gynecologic Oncology*, 166(3), 596–605.  
<https://doi.org/10.1016/j.ygyno.2022.07.024>
- Srivastava, S., Kumar, P., Chaudhry, V., & Singh, A. (2020). Detection of Ovarian Cyst in Ultrasound Images Using Fine-Tuned VGG-16 Deep Learning Network. *SN Computer Science*, 1(2), 81. <https://doi.org/10.1007/s42979-020-0109-6>
- Subramanian, M., Cho, J., Sathishkumar, V. E., & Naren, O. S. (2023). Multiple Types of Cancer Classification Using CT/MRI Images Based on Learning Without Forgetting Powered Deep Learning Models. *IEEE Access*, 11, 10336–10354.  
<https://doi.org/10.1109/ACCESS.2023.3240443>
- Suha, S. A., & Islam, M. N. (2022). An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image. *Scientific Reports*, 12(1), 17123.  
<https://doi.org/10.1038/s41598-022-21724-0>
- Swapnarekha, H., Dash, P. B., Nayak, J., & Routray, A. R. (2023). An Optimistic Bayesian Optimization Based Extreme Learning Machine for Poly-cystic Ovary Syndrome Diagnosis. In J. Nayak, A. K. Das, B. Naik, S. K. Meher, & S. Brahnam (Eds.), *Nature-Inspired Optimization Methodologies in Biomedical and Healthcare* (pp. 175–193). Springer International Publishing.  
[https://doi.org/10.1007/978-3-031-17544-2\\_8](https://doi.org/10.1007/978-3-031-17544-2_8)
- Taleb, N., Mehmood, S., Zubair, M., Naseer, I., Mago, B., & Nasir, M. U. (2022). Ovary Cancer Diag-nosing Empowered with Machine Learning. 2022 International Conference on Business Analytics for Technology and Security (ICBATS), 1–6.  
<https://doi.org/10.1109/ICBATS54253.2022.9759010>
- Tanaka, Y. O., Okada, S., Satoh, T., Matsumoto, K., Oki, A., Saida, T., Yoshikawa, H., & Minami, M. (2016). Differentiation of epithelial ovarian cancer subtypes by use of imaging and clinical data: a detailed analysis. *Cancer Imaging: The Official Publication of the International Cancer Imaging Society*, 16, 3.  
<https://doi.org/10.1186/s40644-016-0061-9>
- Tiwari, S., Kane, L., Koundal, D., Jain, A., Alhudhaif, A., Polat, K., Zaguia, A., Alenezi, F., & Althubiti, S. A. (2022). SPOSDS: A smart Polycystic Ovary Syndrome diagnostic system using machine learning. *Expert Systems with Applications*, 203, 117592.  
<https://doi.org/10.1016/j.eswa.2022.117592>
- Tripathi, A., Arora, A., & Bhan, A. (2021). Classification of cervical cancer using Deep Learning Algorithm. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 1210–1218.  
<https://doi.org/10.1109/ICICCS51141.2021.9432382>
- Visalaxi, S., Punnoose, D., & Muthu, T. S. (2021). An Analogy of Endometriosis Recognition Using Machine Learning Techniques. 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Net-works (ICICV), 739–746.  
<https://doi.org/10.1109/ICICV50876.2021.9388403>
- Wang, G., Zhan, H., Luo, T., Kang, B., Li, X., Xi, G., Liu, Z., & Zhuo, S. (2023). Automated Ovarian Cancer Identification Using End-to-End Deep Learning and Second Harmonic Generation Imaging. *IEEE Journal of Selected Topics in Quantum Electronics*, 29(4: Biophotonics), 1–9.  
<https://doi.org/10.1109/JSTQE.2022.3228567>
- Wang, R., Cai, Y., Lee, I. K., Hu, R., Purkayastha, S., Pan, I., Yi, T., Tran, T. M. L., Lu, S., Liu, T., Chang, K., Huang, R. Y., Zhang, P. J., Zhang, Z., Xiao, E., Wu, J., & Bai, H. X. (2021). Evaluation of a convolutional neural network for ovarian tumor differentiation based on magnetic resonance im-aging. *European Radiology*, 31(7), 4960–4971. <https://doi.org/10.1007/s00330-020-07266-x>
- Wang, R., Gu, Z., Wang, Y., Yin, X., Liu, W., Chen, W., Huang, Y., Wu, J., Yang, S., Feng, L., Zhou, L., Li, L., Di, W., Pu, X., Huang, L., & Qian, K. (2022). A “One-Stop Shop” Decision Tree for Diagnosing and Phenotyping Polycystic Ovarian

Syndrome on Serum Metabolic Fingerprints. *Advanced Functional Materials*, 32(45), 2206670. <https://doi.org/10.1002/ADFM.202206670>

Zahras, D., & Rustam, Z. (2018). Cervical Cancer Risk Classification Based on Deep Convolutional Neural Network. 2018 International Conference on Applied Information Technology and Innovation (ICAITI), 149–153. <https://doi.org/10.1109/ICAITI.2018.8686767>

Zhang, L., Huang, J., & Liu, L. (2023). Retraction Note: Improved Deep Learning Network Based in Combination with Cost-sensitive Learning for Early Detection of Ovarian Cancer in Color Ultrasound Detecting System. *Journal of Medical Systems*, 47(1), 20. <https://doi.org/10.1007/s10916-023-01915-6>

Zhang, Z., & Han, Y. (2020). Detection of Ovarian Tumors in Obstetric Ultrasound Imaging Using Logistic Regression Classifier with an Advanced Machine Learning Approach. *IEEE Access*, 8, 44999–45008. <https://doi.org/10.1109/ACCESS.2020.2977962>

Zhou, H., Wang, S., Zhang, T., Liu, D., & Yang, K. (2021). Ultrasound image analysis technology under deep belief networks in evaluation on the effects of diagnosis and chemotherapy of cervical cancer. *The Journal of Supercomputing*, 77(4), 4151–4171. <https://doi.org/10.1007/s11227-020-03421-9>

Zhou, J., Zeng, Z. Y., & Li, L. (2020). Progress of Artificial Intelligence in Gynecological Malignant Tumors. *Cancer Management and Research*, 12, 12823–12840. <https://doi.org/10.2147/CMAR.S279990>.

## AUTHOR BIOGRAPHIES



Prof. Jyoti Ishwar Nandalwar is an Assistant Professor in the Department of Computer Science and Engineering at Shree Siddheshwar Women's College of Engineering, Solapur. She completed her Bachelor of Engineering (BE) in Computer Science

and Engineering from JDIET, Yavatmal, Maharashtra, in 2003 and her M.Tech from Jawaharlal Nehru Technological University (JNTU), Hyderabad, in 2014. Currently, she is pursuing her PhD from Amravati University, Maharashtra. With over 18.5 years of teaching experience in reputed educational institutions, Prof. Nandalwar has published 14 research articles in reputed national and international journals, 1 research article in a Scopus Indexed Journal, holds 8 patents, and has authored 2 books. She has also participated in 22 workshops/FDPs and organized 20 seminars/workshops/FDPs for Faculties and Students. Her areas of academic interest include Operating Systems, Data Structures, and Computer Networks.



Dr. Pradip Jawandhiya is an accomplished academican and administrator with over 20 years of experience in education and engineering. He is the Principal of P.R. Pote Patil College of Engineering and Management, Amravati. Dr. Jawandhiya holds a Ph.D. in Computer Science & Engineering from Sant Gadge Baba Amravati University, specializing in Mobile Ad hoc Networks. He also earned an MBA in Human Resource Management, along with a B.E. and M.E. in Computer Science & Engineering. Before his current role, he served as Professor and Principal at Pankaj Laddhad Institute of Technology and Management Studies, Buldana, for 11 years. He has also held leadership roles at institutions like Jagdambha College of Engineering & Technology and Jawaharlal Darda Institute of Engineering & Technology, contributing to academic and administrative excellence. Dr. Jawandhiya is a Senior Member of IEEE (SMIEEE) and has served on the Board of Studies for Information Technology at Sant Gadge Baba Amravati University. He is skilled in department administration, computer science, and organizational development. With a strong academic foundation and extensive leadership experience, Dr. Jawandhiya continues to drive innovation in education and technology.

# **Systematic modernization of fish smoking method with the implementation of smoked fish machine based on Internet of Things technology**

Rizal Justian Setiawan<sup>1,6\*</sup>, Khakam Ma'ruf<sup>2</sup>, Darmono<sup>3</sup>, Nur Azizah<sup>4</sup>, Nur Evirda Khosyati<sup>5</sup>

<sup>1</sup> Department of Industrial Engineering and Management, College of Engineering, Yuan Ze University, Zhongli, Republic of China

<sup>2</sup> Department of Industrial Engineering, Faculty of Engineering, Gadjah Mada University, Yogyakarta, Republic of Indonesia

<sup>3</sup> Department of Civil Engineering and Planning Education, Faculty of Engineering, Yogyakarta State University, Yogyakarta, Republic of Indonesia

<sup>4</sup> Department of International Public Health, School of Public Health, China Medical University, Taichung, Republic of China

<sup>5</sup> Department of Culinary Technology Education, Faculty of Engineering, Yogyakarta State University, Yogyakarta, Republic of Indonesia

<sup>6</sup> Department of Marketing Management, Faculty of Economics, Indonesia Open University, Jakarta, Republic of Indonesia

\* Corresponding author E-mail: rizaljustiansetiawan99@gmail.com

(Received .....; Final version received .....; Accepted .....)

## **Abstract**

The traditional process of smoking fish, which is widely used in coastal regions, poses significant challenges due to its labor-intensive nature, constant supervision, and difficulty maintaining stable temperatures. These issues often result in inefficiencies, inconsistent product quality, and potential safety hazards. Given the importance of the smoked fish industry in sustaining local economies in the coastal area of Indonesia, there is a critical need for more advanced, reliable, and efficient methods of fish smoking. This study addresses these challenges by developing an Internet of Things (IoT)-integrated monitoring and control system for the smoked fish machine. This study was conducted to develop a monitoring and control system for machines which includes turning on/off machine components, temperature monitoring, and blower revolutions per minute (RPM) control. The results of the study showed that the implementation of IoT can activate machine components such as a blower, a servo motor, and a light. Moreover, IoT can monitor the machine temperature from a smartphone in real-time by integrating a temperature sensor. The temperature difference between the sensor and the analog thermometer was found to be 0.1 – 0.5°C, proving that the temperature on the IoT system is not very different from the analog thermometer. Furthermore, the blower RPM control results showed that the system could maintain the temperature in the optimal range (75 – 90°C) for smoking the fish, with a maximum deviation of 1°C, and the blower RPM can be adjusted through the control in the IoT system. In general, the use of IoT can simplify machine operation for users.

*Keywords:* Control, Fish, Internet of Things, Machine, Monitoring, Temperature



## 1. Introduction

Technological advancements have led to significant innovations that have profoundly impacted human life (Sánchez & Hartlieb, 2020). The rapid pace of technological progress has resulted in widespread smartphone ownership (Bauer et al., 2020; Leshner, 2021; Senjam et al., 2021; Tapiero et al., 2020), which in turn supports the development of digital technologies such as the Internet of Things (IoT) (Khanh et al., 2022; Mahmood, 2021; Sami et al., 2020). The IoT gained popularity in 2010 by introducing the “smart home” concept and Nest’s launch of a smart thermostat controllable via mobile phones (Verma et al., 2021). Today, nearly everyone possesses a smartphone connected to the internet, facilitating the IoT concept by allowing devices to connect to the internet and communicate autonomously, without human intervention (Baiyere et al., 2020; Karthick & Pankajavalli, 2020). These devices range from household appliances and vehicles to industrial machinery and even simple objects equipped with sensors capable of sending and receiving data (Chen & Setiawan, 2023; Laghari et al., 2021).

The primary function of IoT is to enhance the ease of human life (Setiawan et al., 2021). Through IoT, various devices and systems can be integrated and remotely controlled, offering increased efficiency and convenience (Salih et al., 2022; Shafique et al., 2020; Shammar & Zahary, 2020). Users can remotely monitor and control devices such as temperature regulators, irrigation systems, and security apparatus through IoT-enabled mobile devices (García et al., 2020; Quy et al., 2022; Yasin et al., 2021). IoT facilitates the practical use and control of machines, enabling remote operation and monitoring, thereby reducing direct human involvement and minimizing potential hazards (Kim et al., 2020; Molaei et al., 2020; Wójcicki et al., 2022).

The smoked fish industry, which sustains the livelihoods of many in Indonesia’s coastal areas, is an example of an industry that could benefit from IoT technology (Kruijssen, 2020; Ng et al., 2022; Suwandhahannadi et al., 2024). Currently, most smoked fish enterprises in Indonesia rely on traditional fish smoking methods using stoves (Haryanto et al., 2021). However, these traditional methods are often less effective, efficient, and hygienic compared to modern, controlled solutions. Traditional fish smoking requires continuous supervision and takes a longer time to complete, with unstable temperatures often

leading to product damage (Dash et al., 2022). With technological advancements, more practical smoked fish machines have been developed. Research by Al Hudha et al. (2018) indicates that these machines can be integrated with IoT technology to enhance user convenience. IoT-enabled smoked fish machines can be remotely controlled and monitored, allowing operators to oversee the smoking process in real-time (Al Hudha et al., 2018). The technological development of this sector is crucial for Indonesia, a maritime nation with abundant fishery resources (Setiawan et al., 2023).

The integration and implementation of IoT into smoked fish machines are expected to streamline operations, allowing users to monitor and control the machines in real-time via mobile devices. IoT can save time and energy, as users can perform necessary monitoring and control without being physically present at the site. Additionally, data generated by IoT-enabled smoked fish machines can be analyzed to enhance productivity and efficiency in the production process.

## 2. Literature Review

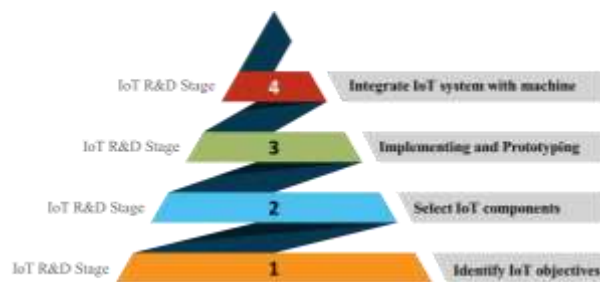
The first study, conducted by Rahman et al. (2020), implemented a temperature monitoring system using DHT11 sensors within an IoT framework. The system employed an Arduino, an ethernet shield, and a DHT11 sensor to transmit data to the AT&T M2X IoT platform via a RESTful API. The M2X platform’s trigger feature was integrated with the IFTTT service, enabling the system to send notifications via email, SMS, and mobile push notifications when extreme changes in temperature and humidity were detected. Data analysis revealed significant temperature and humidity variations across different locations and a strong negative correlation between temperature and humidity, highlighting the need for improved data center layout design (Rahman et al., 2020).

In the second study, Jayanti et al. (2019) developed an oven for smoking fish using the cold smoking method, which incorporated an Arduino Uno-based monitoring system. This study introduced a temperature monitoring system for a smoked fish machine capable of tracking the temperature within the smoke chamber. The Arduino Uno microcontroller connected a temperature sensor, which could accurately read the internal temperature, and an LCD that displayed real-time temperature conditions (Jayanti et al., 2019).

The third study involved the development of a smoked fish machine with an IoT-based control system by Al Hudha et al. (2019). The machine utilized an IoT system to activate machine components and monitor temperature, thereby enhancing user access to the machine's features. Based on the findings from these three studies—each of which explored the application of the DHT11 temperature sensor, Arduino Uno microcontroller, and IoT in control and monitoring systems—it is evident that these components can be effectively and efficiently implemented. Consequently, integrating these technologies into a smoked fish machine is both feasible and advantageous.

### 3. Research Methodology

This study employed a research and development (R&D) methodology and utilized the stages of IoT system development as delineated by Dilmegani (2024), as illustrated in Fig. 1.



**Fig. 1.** IoT research and development Stages (Dilmegani, 2024).

The research process encompasses four distinct phases of IoT development, which are elaborated upon as follows:

#### 3.1. Identify Internet of Things Objectives

At this stage, it is crucial to define the primary objectives of IoT system development clearly. Generally, the application of IoT in smoked fish machines aims to provide a user-friendly interface that facilitates remote monitoring and control of the smoking room. The specific objectives to be achieved include:

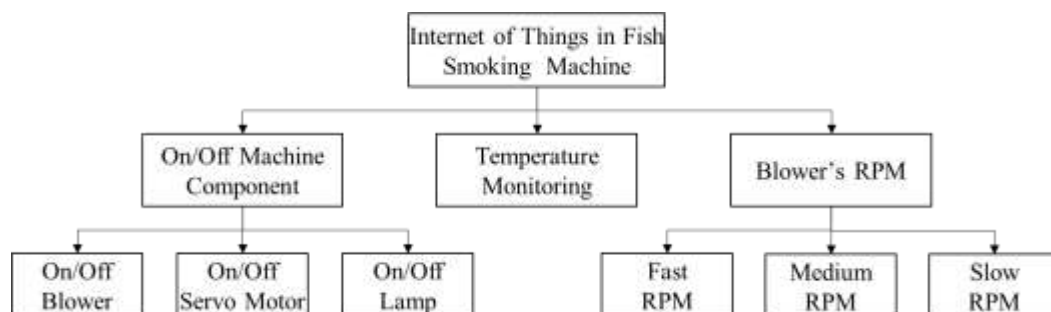
- Controlling the activation and deactivation of the blower, servo motor for driving the machine's shaft, and the machine's lighting.
- Real-time monitoring of the temperature within the fish smoking room (inside of the smoking tube).
- Regulating the heat in the combustion chamber by adjusting the blower's revolutions per minute (RPM).

As depicted in the flowchart in Fig. 2, these objectives represent the visualized goals of the IoT system that will be implemented in the smoked fish machine for this study. Three objectives include machine components, temperature monitoring, and the blower's RPM.

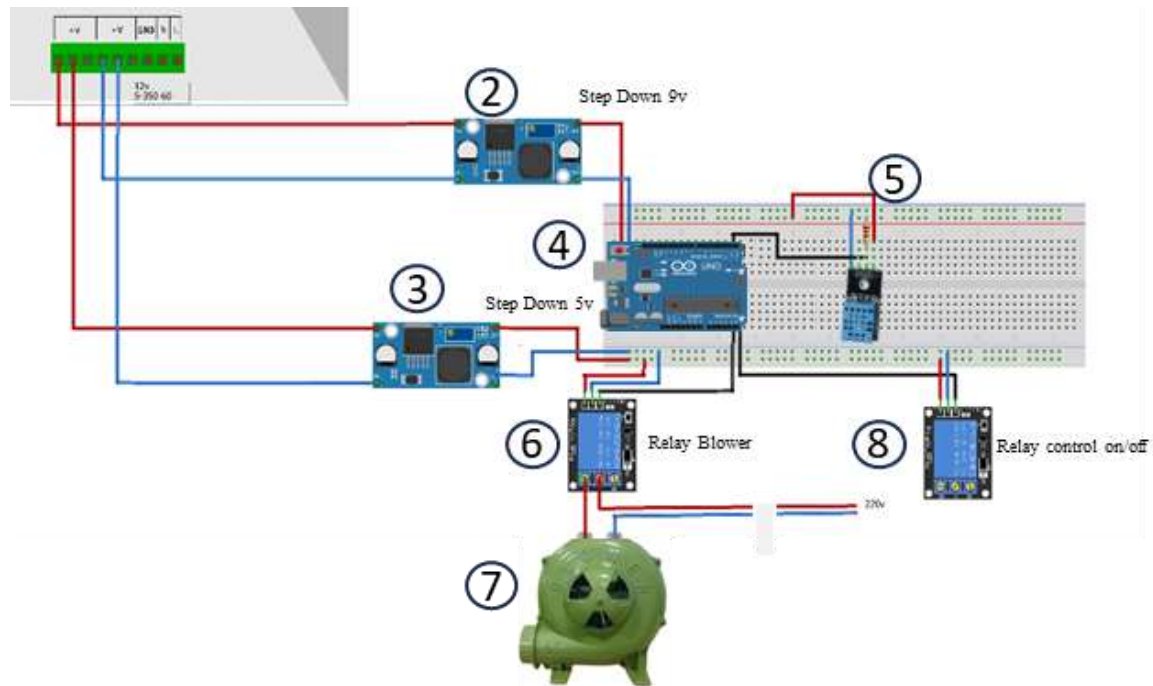
#### 3.2. Selection of IoT Component

Once the objectives for the IoT system were established, the next crucial step involved selecting the necessary components. This research identifies six key components essential for preparing the IoT system:

- The microcontroller serves as the central processing unit of the IoT system, responsible for processing data from sensors and controlling other connected components.
- The DHT11 temperature sensor is employed to accurately and continuously measure the temperature and humidity within the fish smoking chamber in real-time.



**Fig. 2.** Internet of Things objectives in this study.



**Fig. 3.** Internet of Things electrical system design plan.

- c) The blower's electronic system, which manages air circulation within the smoking chamber, is automatically integrated with the IoT system for optimal operation.
- d) A WiFi or Ethernet module is required to connect the IoT system to the internet, enabling remote data transmission and reception.
- e) The user interface on a mobile smartphone provides the platform for users to interact with and control the IoT system.

After identifying the primary components for the IoT system in the fish smoking chamber, a comprehensive design was developed, incorporating both essential and supporting components. The electronic circuit design, as illustrated in Fig. 3, presents the complete schematic of the IoT system, detailing the integration of various components, including sensors, microcontrollers, communication modules, and output devices.

The detailed schematic of the IoT components depicted in Fig. 3 is further elaborated in Table 1, which comprehensively explains their specifications.

### 3.3. Implementation and Prototyping

Following the selection of IoT components and the completion of the system design, the next stage involves the implementation and prototyping process.

This stage is divided into two distinct activities: hardware prototype implementation and software prototype implementation.

**Table 1.** Internet of Things electronic system components.

No	Detail part
1.	Power supply 12 V Microcontroller
2.	Modul step down 9 V
3.	Modul step down 9 V
4.	Microcontroller Arduino Uno
5.	Temperature Sensor DHT 11
6.	Relay setting RPM Blower
7.	Blower 2" 220 V
8.	Relay on/off Machine

#### a) Hardware Prototype Implementation

All IoT components are assembled and integrated for hardware prototyping implementation according to the finalized design specifications. The process begins with connecting the DHT11 temperature sensor to the Arduino Uno microcontroller. Subsequently, program code is developed to acquire real-time temperature data from the sensor and transmit it to the user interface via an internet connection. Additionally, the blower RPM and timer system are integrated with the microcontroller, with a corresponding program code designed to

regulate the blower RPM automatically based on the target temperature. The system also includes programming for the on/off activation of the blower, servo motor, and additional components.

A WiFi or Ethernet module is integrated into the microcontroller to facilitate internet connectivity for the IoT system. User interfaces, such as web or mobile applications, are then developed to enable remote monitoring of temperature and humidity and control of temperature settings, fan speeds, and timers. Upon integration of all components, the IoT system prototype undergoes rigorous testing. This testing involves simulating various temperature conditions and smoking scenarios to ensure that the system operates as intended and meets all functional requirements.



**Fig. 4.** Box panel and Internet of Things component assembly.

Fig. 4 illustrates the design of the control printed circuit board circuit, which is intended to be integrated into the IoT system for monitoring and controlling the fish-smoking process. This electrical system is installed on the machine's control panel, with careful attention to both functionality and safety. The design includes several key components: a power supply, a step-down module, a DHT11 temperature sensor, a timer sensor, an Arduino Uno microcontroller, an IoT communication module, the machine's heating system, and the Blynk IoT platform. The DHT11 temperature sensor is integral to the system, providing real-time temperature measurements during the smoking process. Its selection was based on its accuracy in measuring both temperature and humidity, as well as its capability for digital communication with the microcontroller. The Arduino Uno microcontroller facilitates wireless communication via a Wi-Fi network, enabling seamless integration with the IoT system.

## b) Software Prototype Implementation

Fig. 5 illustrates the coding process for the oven temperature monitoring system, which employs the DHT11 sensor in conjunction with an Arduino Uno microcontroller and the Blynk platform.

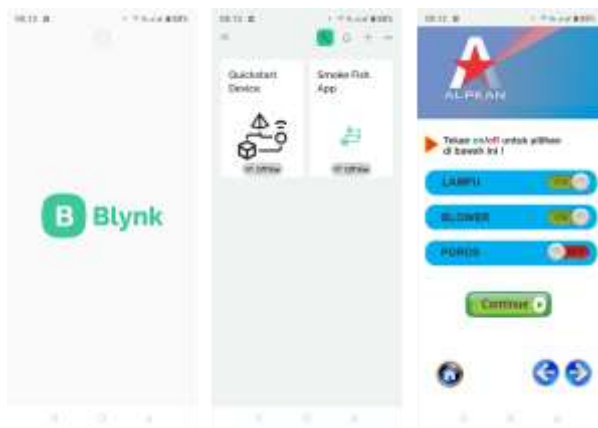


**Fig. 5.** Internet of Things coding on smoked fish machine

The program is designed to read temperature data from the DHT11 sensor, display it on an LCD screen, and transmit it to the Blynk application for real-time monitoring. Blynk facilitates the display of both current temperature and local time, with the program also incorporating functionality to retrieve local time using the TimeZone API.

The user interface is designed to enable interaction with the system through a custom application. This application allows users to monitor the smoking process temperature and regulate the blower's RPM and speed control. Access to the monitoring system requires users to log in with credentials previously created. The integration of the IoT system with the Blynk application is depicted in Fig. 6, which provides a visual representation of the system's status and functionality.

Fig. 6 illustrates the design results of the IoT monitoring and control system utilizing the Blynk platform. This stage is critical for the effective implementation of the system. The Blynk platform facilitates real-time data transmission from the temperature sensor to the Blynk cloud server via the Arduino Uno IoT communication module. Subsequently, this data is accessible and visualizable through both the Blynk mobile application and web interface. Blynk's user interface offers an intuitive and user-friendly display for real-time sensor monitoring, including features such as historical temperature graphs and notification alerts for parameter deviations.



**Fig. 6.** Internet of Things user interface design display.

Additionally, the Blynk interface enables remote control of the blower's RPM through virtual buttons and allows users to set target temperatures. The system's IoT connectivity relies on a stable Wi-Fi network. Connectivity performance was assessed by monitoring the reliability of data transfer from the microcontroller to the Blynk platform. Security considerations are integral to the system's design, with the Blynk platform providing robust authentication and data encryption to safeguard communications between devices and cloud servers. Access to the Blynk user interface is secured with password protection, ensuring that only authorized users can monitor and control the system. Testing results confirm that data transmission is stable and efficient, with minimal or no significant data loss.

### 3.4. Integrate Internet of Things System with Smoked Fish Machine

Following the successful implementation of the IoT system prototype for the smoking machine, the subsequent phase involved integrating the system with the operational machine. This study required the IoT system to be interfaced with both the smoked fish machine and a smartphone. The IoT system was designed for remote monitoring and control, provided that the machine is connected to a nearby Wi-Fi router and the user's smartphone has an active internet connection. In scenarios where Wi-Fi is unavailable, users can employ an alternative IoT connectivity solution by enabling tethering on a smartphone, which connects to the IoT panel on the machine within a maximum distance of 30 meters.

Fig. 7 shows the integration process between the IoT system and the smoked fish machine components. This stage is crucial for ensuring that the IoT application on the smartphone and the IoT panel on the smoking machine function cohesively. Calibration procedures were also performed at this stage to verify the validity and accuracy of data measurements.



**Fig. 7.** Integrating the Internet of Things with a smoking fish machine.

## 4. Result and Discussion

### 4.1. Internet of Things Implementation to Turn On / Off Machine Components

The flow diagram depicted in Fig. 8 illustrates the operational sequence of the IoT system applied to a smoked fish machine for component activation. The process initiates with the manual powering on of the machine, followed by the application of power through a supply unit that converts 220 V to 12 V. This is further regulated by a step-down converter that reduces the voltage to 9 V, ensuring that the current is within safe operational limits for the Arduino Uno microcontroller. Once the machine is successfully powered on, the user must engage in the IoT mode by pressing the designated IoT button on the smoked fish machine. The user can control the machine remotely by establishing a connection between the machine and the smartphone via the IoT system. The IoT system facilitates the activation and deactivation of three components: the blower, servo motor, and light.



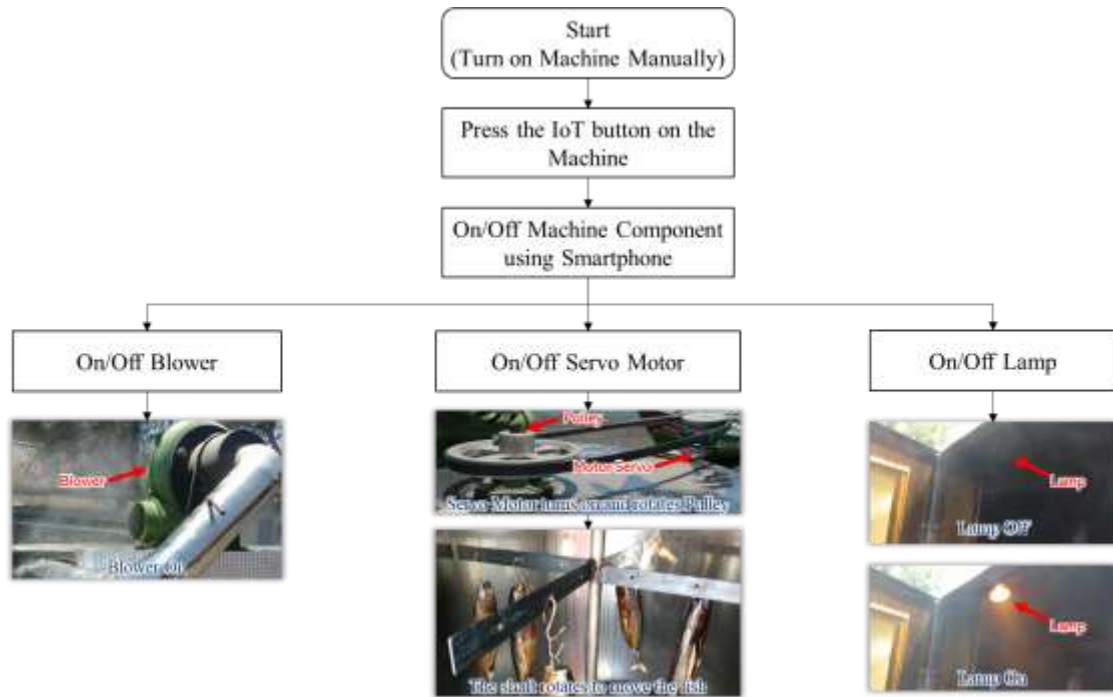


Fig. 8. Internet of Things system to activate components in the smoked fish machine.

#### 4.2. Internet of Things Performance Testing on Temperature Monitoring

The microcontroller plays a pivotal role in the IoT system by collecting data from the DHT11 temperature sensor, processing this data, and transmitting it to the Blynk IoT platform via the communication module. The Blynk application is utilized to monitor the temperature data collected. The microcontroller also regulates the blower's RPM to maintain an optimal level for effective smoking. Additionally, a timer system is integrated to cease smoking once the designated time has elapsed automatically. Following the installation of the IoT system on the smoking machine, a thorough evaluation was conducted to assess the performance of the RPM blower monitoring and control system during the fish-smoking process. Initial tests focused on evaluating the accuracy of the DHT11 temperature sensor in monitoring the parameters of the smoking process.

Table 2 presents the test results, demonstrating that the DHT11 temperature sensor exhibits high accuracy, with an average deviation of 0.1–0.5°C compared to the analog thermometer used for comparison. This level of accuracy is crucial for ensuring that the temperature data obtained accurately reflects the actual conditions within the smoking machine.

Table 2. Temperature monitoring performance testing.

No	Machine temperature (°C)		Status machine/gas solenoid mode
	IoT application	Analog thermometer	
1	34.1	34.4	off
2	43.5	43.3	on / Fast RPM Blower
3	48.8	48.6	on / Fast RPM Blower
4	58.7	58.2	on / Fast RPM Blower
5	62.6	62.5	on / Fast RPM Blower
6	69.2	69	on / Fast RPM Blower
7	73.8	73.6	on / Fast RPM Blower
8	79.3	79	on / Medium RPM Blower
9	85.3	85.2	on / Medium RPM Blower
10	94.6	94.5	on / Slow RPM Blower
11	100.7	100.5	on / Slow RPM Blower

#### 4.3 Internet of Things Implementation to Control Blower's Revolutions Per Minute

A blower control system, based on RPM, is employed to regulate hot air circulation during the fish-smoking process. This system integrates a DHT11 temperature sensor with IoT technology and a microcontroller to monitor the temperature within the dryer continuously. The effectiveness of this monitoring system in maintaining optimal temperature levels during smoking was assessed through testing. The blower system is calibrated to sustain a temperature range of 75 – 90°C. Test results indicate

that the RPM control system effectively maintains the temperature within the specified range, with a maximum deviation of 1°C. Additionally, the system demonstrated prompt responsiveness to changes in the smoking conditions, adjusting the blower RPM rapidly in response to significant temperature fluctuations within the smoking chamber. This real-time parameter control enhances the efficiency and effectiveness of the smoking process.

Fig. 9 illustrates the components installed for controlling the blower's RPM, including a dimmer silicon-controlled rectifier and a mini servo. This system integrates a dimmer and a servo motor to regulate power settings automatically. The servo motor adjusts the dimmer lever, enabling automated RPM control without manual intervention, thereby enhancing blower efficiency.



Fig. 9. Blower RPM control system installation.

The RPM control system modulates air circulation in the smoking machine based on the temperature within the drying chamber. When the temperature exceeds 100°C, the RPM is reduced to

500 to prevent overheating, which could damage the fish. At temperatures between 75 – 90°C, considered optimal for smoking, the blower operates at a medium speed of 1,500 RPM to maintain the ideal conditions. If the temperature falls below 75°C, the system increases the blower RPM to 2,500, facilitating rapid replacement of the air with hotter smoke to raise the temperature in the smoking chamber effectively.

Fig. 10 depicts the operational scheme of the RPM blower control system. This automated control system is designed to stabilize and optimize the heat generated from combustion smoke, ensuring that the temperature within the smoking processing tube remains consistently within the target range of 75 – 90°C. The Blower RPM Control System is integrated with the IoT framework to regulate temperature fluctuations throughout the smoking process. This integration allows for precise adjustments to the temperature, either increasing or decreasing it as necessary.

The system's design aims to maintain the temperature in the smoking processing tube efficiently, adhering to the optimal range of 75 – 90°C as established in the literature (Tahir et al., 2020). Automating the temperature regulation process ensures that the fish smoking procedure operates at peak efficiency and effectiveness.

## 5. Conclusion

This study successfully developed and implemented an IoT-integrated monitoring and control system for a smoked fish machine. The IoT system allows for precise control of the machine's

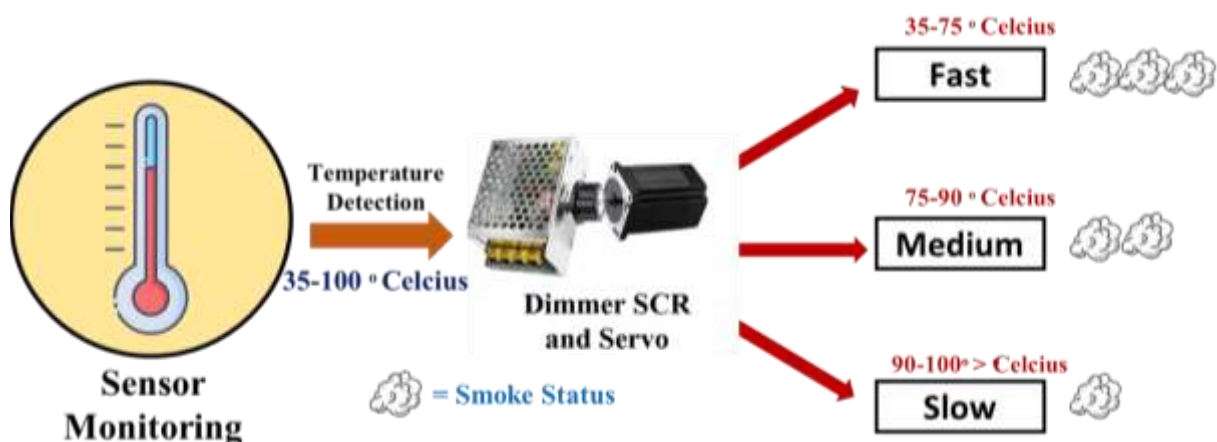


Fig. 10. Working scheme for silicon-controlled rectifier dimmer and revolutions per minute blower control system design.



components, enabling users to turn them on or off remotely. Additionally, the system facilitates real-time temperature monitoring within the fish smoking room. It also regulates the heat in the combustion chamber by adjusting the blower's RPM.

Implementing the IoT system to the machine offers users significant flexibility in monitoring and controlling the fish-smoking process from remote locations, eliminating the need for physical presence at the machine's site. Performance testing revealed that the IoT system operates optimally, with only a brief delay of approximately one second when activating machine components, such as a blower, servo motor, and light, via a smartphone interface.

Moreover, testing demonstrated that the IoT-connected DHT11 sensor accurately measures temperature and humidity, with an average discrepancy of just 0.1 – 0.5°C compared to an analog thermometer. The blower RPM control system was also found to be effective in maintaining the temperature within the optimal range, with a maximum deviation of 1°C, while efficiently adjusting hot air circulation based on the temperature conditions.

Overall, the results of this study indicate that the application of IoT technology to smoked fish machines has considerable potential for enhancing operational efficiency and control. Future development efforts could focus on integrating additional features, such as advanced data analysis and predictive maintenance systems, to optimize the system's performance further.

## References

- Al Hudha, M. E., Setiawan, R. J., & Fauzi, I. (2018). Smofim: Mesin Pengasap Ikan Berbasis Solar Photovoltaic Terintegrasi Android Mobile Iot (Internet of Things) Dengan Exhaust Filter Pereduksi Polutan Co, Co2 Dan Hc Sebagai Upaya Untuk Meningkatkan Perekonomian Masyarakat Nelayan Di Pantai Trisik Kulon Progo. *Jurnal Ilmiah Penalaran dan Penelitian Mahasiswa*, 2(1), 42-51.
- Baiyere, A., Topi, H., Venkatesh, V., & Donnellan, B. (2020). The internet of things (IoT): A research agenda for information systems. *Communications of the Association for Information Systems*, 47.
- Bauer, M., Glenn, T., Geddes, J., Gitlin, M., Grof, P., Kessing, L. V., ... & Whybrow, P. C. (2020). Smartphones in mental health: a critical review of background issues, current status and future concerns. *International Journal of Bipolar Disorders*, 8, 1-19.
- Chen, Y.T., & Setiawan, R.J. (2023). Energy Saving Solution for Welding Process: A SME Case Study of Fume Extractor. *2023 IEEE 6th International Conference on Knowledge Innovation and Invention (ICKII)*, Sapporo, Japan.
- Dash, K. K., Fayaz, U., Dar, A. H., Shams, R., Manzoor, S., Sundarsingh, A., ... & Khan, S. A. (2022). A comprehensive review on heat treatments and related impact on the quality and microbial safety of milk and milk-based products. *Food Chemistry Advances*, 1, 100041.
- Dilmegani, C. (2024). IoT Implementation: Steps & Best Practices in 2024. *Ai Multiple Research*.
- García, L., Parra, L., Jimenez, J. M., Lloret, J., & Lorenz, P. (2020). IoT-based smart irrigation systems: An overview on the recent trends on sensors and IoT systems for irrigation in precision agriculture. *Sensors*, 20(4), 1042.
- Haryanto, A., Hidayat, W., Hasanudin, U., Iryani, D. A., Kim, S., Lee, S., & Yoo, J. (2021). Valorization of Indonesian wood wastes through pyrolysis: A review. *Energies*, 14(5), 1407.
- Jayanti, T.A.D., Sudarmanto, A., & Faqih, M.I. (2019). Cold Smoking Equipment Design Of Smoked Fish Products With Closed Circulation Using Temperature and Concentration Monitoring System Based On Arduino Uno. *IOP Conference Series: Materials Science and Engineering*, 846, 012025.
- Karthick, G. S., & Pankajavalli, P. B. (2020). A review on human healthcare internet of things: a technical perspective. *SN Computer Science*, 1(4), 198.
- Khanh, Q. V., Hoai, N. V., Manh, L. D., Le, A. N., & Jeon, G. (2022). Wireless communication technologies for IoT in 5G: Vision, applications, and challenges. *Wireless Communications and Mobile Computing*, 2022(1), 3229294.
- Kim, W. S., Lee, W. S., & Kim, Y. J. (2020). A review of the applications of the internet of things (IoT) for agricultural automation. *Journal of Biosystems Engineering*, 45, 385-400.
- Kruijssen, F., Tedesco, I., Ward, A., Pincus, L., Love, D., & Thorne-Lyman, A. L. (2020). Loss and waste in fish value chains: A review of the evidence from low and middle-income countries. *Global Food Security*, 26, 100434.
- Kusumanti, I., Yulianti, W., & Jannah, N. (2021). Physiochemical property of wastewater discharged from smoked fish industry around fishponds area in Penatarsewu Village, Sidoarjo, East Java. *IOP Conference Series: Earth and Environmental Science*, 744, 012037.
- Laghari, A. A., Wu, K., Laghari, R. A., Ali, M., & Khan, A. A. (2021). A review and state of art of

- Internet of Things (IoT). *Archives of Computational Methods in Engineering*, 29, 1395-1413.
- Leshner, A. P., Gavrilova, Y., Ruggiero, K. J., & Evans, H. L. (2021). Surgery and the smartphone: can technology improve equitable access to surgical care?. *Journal of Surgical Research*, 263, 1-4.
- Mahmood, S. (2021). Review of internet of things in different sectors: recent advances, technologies, and challenges. *Journal on Internet of Things*, 3(1), 19.
- Molaei, F., Rahimi, E., Siavoshi, H., Afrouz, S. G., & Tenorio, V. (2020). A comprehensive review on internet of things (IoT) and its implications in the mining industry. *American Journal of Engineering and Applied Sciences*, 13(3), 499-515.
- Ng, C. K. C., & Ong, R. C. (2022). A review of anthropogenic interaction and impact characteristics of the Sundaic mangroves in Southeast Asia. *Estuarine, Coastal and Shelf Science*, 267, 107759.
- Quy, V. K., Hau, N. V., Anh, D. V., Quy, N. M., Ban, N. T., Lanza, S., ... & Muzirafuti, A. (2022). IoT-enabled smart agriculture: architecture, applications, and challenges. *Applied Sciences*, 12(7), 3396.
- Rahman, R., Hashim, U. R. A., & Ahmad, S. (2020). IoT-based temperature and humidity monitoring framework. *Bulletin of Electrical Engineering and Informatics*, 9(1), 229-237.
- Setiawan, R. J., Tarnadi, A., & Surfani, I. (2021). Design and Manufacture an Automatic Mushroom Sprinkler based Internet of Things to Increase Oyster Mushroom Productivity. *Jurnal Material dan Proses Manufaktur*, 5(1), 1-9.
- Setiawan, R. J., Chen, Y. T., & Suryanto, I. D. (2023). Cost-Effective Fish Storage Device for Artisanal Fishing in Indonesia-Utilization of Solar Cool Box. In *2023 IEEE 17th International Conference on Industrial and Information Systems (ICIIS)* (pp. 471-476). IEEE.
- Salih, K. O. M., Rashid, T. A., Radovanovic, D., & Bacanin, N. (2022). A comprehensive survey on the Internet of Things with the industrial marketplace. *Sensors*, 22(3), 730.
- Sami, N., Mufti, T., Sohail, S. S., Siddiqui, J., Kumar, D., & Neha. (2020). Future Internet of Things (IOT) from Cloud perspective: Aspects, applications and challenges. *Internet of Things (IoT) Concepts and Applications*, 515-532.
- Sánchez, F., & Hartlieb, P. (2020). Innovation in the mining industry: Technological trends and a case study of the challenges of disruptive innovation. *Mining, Metallurgy & Exploration*, 37(5), 1385-1399.
- Senjam, S. S., Manna, S., & Bascaran, C. (2021). Smartphones-based assistive technology: accessibility features and apps for people with visual impairment, and its usage, challenges, and usability testing. *Clinical Optometry*, 311-322.
- Shafique, K., Khawaja, B. A., Sabir, F., Qazi, S., & Mustaqim, M. (2020). Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. *IEEE Access*, 8, 23022-23040.
- Shammar, E. A., & Zahary, A. T. (2020). The Internet of Things (IoT): a survey of techniques, operating systems, and trends. *Library Hi Tech*, 38(1), 5-66.
- Suwandhahannadi, W. K., Le De, L., Wickramasinghe, D., & Dahanayaka, D. D. G. L. (2024). Community participation for assessing and managing ecosystem services of coastal lagoons: A case of the Rekawa Lagoon in Sri Lanka. *Ocean & Coastal Management*, 251, 107069.
- Tahir, M., Salengke, S., Mursalim., Metusalach., & Caesarendra, W. (2020). Performance of Smokehouse Designed for Smoking Fish with the Indirect Method. *Processes*, 8(2), 204.
- Tapiero, S., Yoon, R., Jefferson, F., Sung, J., Limfueco, L., Cottone, C., ... & Clayman, R. V. (2020). Smartphone technology and its applications in urology: a review of the literature. *World Journal of Urology*, 38, 2393-2410.
- Verma, R., Mishra, P. K., Nagar, V., & Mahapatra, S. (2021). Internet of things and smart homes: a review. *Wireless Sensor Networks and the Internet of Things*, 111-128.
- Wójcicki, K., Biegańska, M., Paliwoda, B., & Górna, J. (2022). Internet of things in industry: Research profiling, application, challenges and opportunities—a review. *Energies*, 15(5), 1806.
- Yasin, H. M., Zeebaree, S. R., Sadeeq, M. A., Ameen, S. Y., Ibrahim, I. M., Zebari, R. R., ... & Sallow, A. B. (2021). IoT and ICT based smart water management, monitoring and controlling system: A review. *Asian Journal of Research in Computer Science*, 8(2), 42-56.

## AUTHOR BIOGRAPHIES



**Rizal Justian Setiawan.** He obtained his A.Md.T and B.Ed. in Mechanical Engineering Education from Yogyakarta State University and M.M. in Marketing Management from Universitas Terbuka, Indonesia. Currently, he is a Master's

student in the Industrial Engineering and Management study program at Yuan Ze University, Republic of China.



**Khakam Ma'ruf.** He is a student in the Industrial Engineering study program at Gadjah Mada University, Indonesia. Currently, he is a project manager at a small and medium (SME) manufacturing company named CV. Inanri Sukses Bersama. His main role is to monitor the production and distribution of appropriate technology.



**Darmono.** He is a lecturer at the Faculty of Engineering, Yogyakarta State University. He was appointed a lecturer at the university with a scientific concentration in Civil Engineering and Planning Education. He obtained his education degree such as Drs., Ir., and

Dr. at Yogyakarta State University. Then, he obtained an M.T. degree at Gadjah Mada University.



**Nur Azizah.** She is a researcher from Sriwijaya University. She obtained S.Ked and M.D. degrees at Sriwijaya University. Currently, she is a student in the International Public Health program at China Medical University, Republic of China.



**Nur Evirda Khosyati.** She is an active student in the culinary technology education study program at the Culinary Technology and Fashion Education Department of Yogyakarta State University, Indonesia. She is active in research activity and experimental engineering.

# Knowledge management capability and innovation ambidexterity: The role of intellectual capital and intangible competitive advantage

Nirusa Sirivariskul<sup>1\*</sup>

<sup>1</sup>Department of Management, Faculty of Management Technology,  
Rajamangala University of Technology Isan, Surin Campus, Thailand

\*Corresponding author E-mail: nirusa.si@rmuti.ac.th

(Received 23 August 2024; Final version received 16 November 2024; Accepted 19 November 2024)

## Abstract

This study examines the impact of intellectual capital enabled knowledge capability management on innovation ambidexterity and the moderating role of intangible resources advantage through the lens of open innovation. The sample consists of 105 companies in the Thai food industry. To enhance understanding of how companies can achieve success in innovation, an interesting result of the study is that it helps companies explore new knowledge and leverage existing or new knowledge to innovate more continuously. Based on the presented findings, intangible resources advantage plays a positive moderating role in this relationship: knowledge capability management and intangible resources advantage work together to foster innovation ambidexterity.

**Keywords:** Intangible Resources Advantage, Intellectual Capital, Innovation Ambidexterity, Knowledge Capability Management

## 1. Introduction

Globalization and technological advancements have created intense competition in the business environment. As a result, companies must innovate and introduce new products or services. In today's rapidly evolving business landscape, organizations are increasingly recognizing the pivotal role of knowledge management and innovation ambidexterity in sustaining competitive advantage. The dynamic interplay between these elements can significantly influence an organization's ability to navigate the complexities of modern markets. At the heart of this interplay lies intellectual capital, encompassing the collective knowledge, skills, and experiences within an organization. The strategic management of this intellectual capital cultivates intangible competitive advantages that are crucial for long-term success.

For example, Lei et al. (2020) found that a knowledge-centered culture positively mediates the relationship between HRM practices and innovation

capability, significantly enhancing the impact of HRM practices on knowledge management capability and both exploitative and exploratory innovation. Intellectual capital, which includes human capital, structural capital, and relational capital, plays a critical role in enabling organizations to effectively manage knowledge and foster innovation. Human capital refers to the experience, education, skills, and distinctive attributes of an organization's employees (Cuganesan, 2005); structural capital refers to the procedures, systems, internal structures, and organizational culture (Lee, 2011); and relational capital refers to connections with stakeholders, including suppliers, clients, governments, and society at large. It also encompasses additional relational resources such as the firm's image and customer loyalty (Machorro et al., 2016). Together, these components form a robust foundation for knowledge management and innovation.

The concept of intangible competitive advantage, rooted in the strategic utilization of intellectual capital, emphasizes the importance of non-physical assets that provide a sustainable edge over competitors. Unlike tangible assets such as machinery and buildings, intangible assets are often harder to replicate, thereby offering a distinct source of competitive differentiation. When these intangible assets are effectively leveraged through knowledge management practices, they can drive both exploratory and exploitative innovation, thereby enhancing an organization's ambidexterity. Previous studies, such as those by Hsu and Sabherwal (2011), have examined how knowledge management mediates the impact of intellectual capital on innovation. However, most research on the interaction between intellectual capital, knowledge management strategies, and innovation has focused on industrialized countries. The relative scarcity of empirical data from emerging economies, particularly Thailand, prompted this study. This paper aims to explore the intricate relationships between knowledge management capability, innovation ambidexterity, intellectual capital, and intangible competitive advantage. By examining these interconnections, we seek to provide insights into how organizations can strategically harness their intellectual capital to foster innovation, manage knowledge effectively, and achieve a sustainable competitive advantage in the contemporary business environment.

This work makes several contributions to the field of innovation. The research objectives of this study are:

- To provide new evidence on how intellectual capital enables the management of knowledge capability to increase innovation ambidexterity in prior intellectual capital research.
- To develop this argument by focusing on knowledge capability management in the food industry.

This research follows these steps:

1. Literature review: The initial step involves reviewing literature to analyze the alignment and relationships of conceptual frameworks.
2. Theoretical explanation: The second step entails explaining theories to lay the foundation and develop research hypotheses.

3. Methodological description: The third step explains the research methodology, including sample groups, data, and measurement techniques.
4. Research analysis: The fourth step involves analyzing the research findings.
5. Discussion and conclusion: The final step includes discussing the results and summarizing the research findings.

## 2. Literature Review

### 2.1. Intellectual Capital and Knowledge Management Capability

The relationship between intellectual capital and knowledge management capability is fundamentally synergistic. Intellectual capital, which includes human capital (skills, knowledge, and expertise of employees), structural capital (processes, databases, and organizational routines), and relational capital (networks and relationships with stakeholders), provides the essential resources for knowledge creation and utilization. These three forms of intellectual capital are critical to both company performance (Ahmed et al., 2019; Pedro et al., 2018) and innovation (Allameh, 2018). Knowledge management capability leverages these intellectual assets by systematically acquiring, organizing, sharing, and applying knowledge within the organization. This capability ensures that the valuable insights and expertise embedded in intellectual capital are effectively utilized to enhance decision-making, foster innovation, and improve overall organizational performance. According to Kaufmann and Schneider (2004) and Youndt *et al.* (2004), intellectual capital consists of knowledge and intangible assets that an organization can use to its advantage in order to create economic value and obtain a competitive edge. Thus, a robust knowledge management capability maximizes the potential of intellectual capital, leading to a more knowledgeable, agile, and competitive organization.

### 2.2. Knowledge Management Capability and Innovation Ambidexterity

Knowledge management, defined as the systematic process of creating, sharing, and utilizing knowledge to achieve organizational objectives, serves as the foundation for fostering innovation ambidexterity. Knowledge management capability leverages this capital by systematically acquiring,

sharing, and applying knowledge, thereby enabling the organization to utilize its intellectual assets effectively. This capability is crucial for fostering innovation ambidexterity, the balanced pursuit of both exploratory and exploitative innovations. By effectively managing knowledge, organizations can enhance their ability to innovate continuously, thereby creating a virtuous cycle where successful innovations further enrich intellectual capital, ultimately driving sustained competitive advantage and growth.

While exploitation refers to knowledge obtained through the selection, improvement, and repurposing of current routines built upon consolidated knowledge bases, exploration entails a shift toward new knowledge pathways (Messeni Petruzzelli, 2019). The process of exploration involves creating new knowledge, recombining existing diverse knowledge vectors, or combining old and new knowledge (Carnabuci & Operti, 2013). Ambidextrous businesses are able to explore new opportunities and capitalize on their current competencies with equal dexterity (Lubatkin et al., 2006).

According to Cabrilo and Dahms (2020), innovation ambidexterity can be facilitated by intellectual capital, as it can augment a company's capacity to explore and obtain novel insights and methods that surpass its current expertise. Tseng (2016) suggests that the concept of knowledge management capability (KMC) defines the functional boundaries of an organization's capacity to manipulate knowledge, encompassing the generation, transfer, integration, sharing, and application of information to generate new knowledge. Innovation ambidexterity, which refers to the capability of an organization to simultaneously explore new opportunities (exploratory innovation) and exploit existing capabilities (exploitative innovation), is essential for maintaining a competitive edge in both stable and turbulent environments.

### **2.3. Intangible Competitive Advantage as Moderator for Knowledge Management Capability and Innovation Ambidexterity**

According to Barney (1991) and Kim and Oh (2004), the idea of leveraging an organization's resources to gain a competitive edge suggests that resources such as competencies, assets, capabilities, information, and knowledge can set an organization apart and create unique opportunities. Academics and organizational managers have categorized resources

into five groups: physical, human, technological, financial, and organizational resources (Hofer and Schendel, 1978). According to Penrose's (1959) Resource-Based View of the Firm (RBV), effective resource use is the foundation for organizational expansion. Organizations can use these resources, which include assets, skills, processes, and knowledge, to develop and implement competitive strategies.

According to Hall (1992), intangible resources consist of non-financial elements, whereas tangible resources are composed of financial aspects or physical assets. Intangible Competitive Advantage (ICA) moderates the relationship between Knowledge Management Capability (KMC) and Innovation Ambidexterity (IA) by leveraging non-physical assets such as brand reputation, organizational culture, intellectual property, and unique capabilities. These intangible assets enhance the effectiveness of knowledge management processes by facilitating better knowledge sharing, resource allocation, and strategic alignment. As a result, organizations can more effectively balance and manage both exploratory (radical) and exploitative (incremental) innovation activities, thereby sustaining long-term competitiveness and adaptability in dynamic market environments.

### **2.4. Conceptual Framework and Hypotheses**

Based on the conceptual framework of the research stated previously, the hypotheses of this research were as follows:

**H1:** Intellectual capital was positively associated with knowledge management capability.

**H2:** Intangible resources advantage positively moderated the relationship between knowledge management capability and innovation ambidexterity.

**H3:** knowledge management capability was positively associated with innovation ambidexterity.

## **3. Research Methodology**

### **3.1. Sample**

The online questionnaire was designed to assess the fundamental constructs outlined in the conceptual model. The survey, distributed online, was sent to 395 companies listed in the database of registered food manufacturing firms in Thailand. Utilizing the sampling frame, an email containing a letter detailing the rationale and objectives of the study was sent to the provided email address of each company. Among the

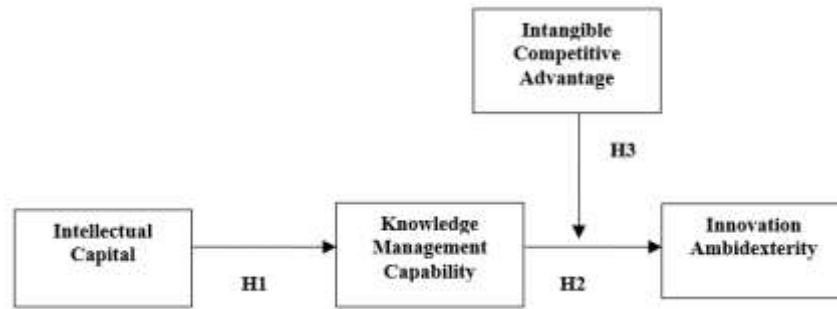


Fig. 1. Research framework.

112 responses initially received, 105 were deemed complete and suitable for analysis related to the variables of interest in our study. The data set gathered encompassed a diverse spectrum of firms in terms of age, size, and type, resulting in an estimated response rate of around 20%.

Furthermore, a test for non-response bias (Armstrong & Overton, 1977) was conducted by comparing demographic characteristics (firm size and firm age) between the earliest and latest responses. Additionally, comparisons between respondent firms and the entire manufacturing firms' population for two available population-level variables firm size and firm age were performed. In addition to the primary constructs outlined in the hypothesis model, this study incorporates control variables recognized to influence firms' competitive advantages: firm size and firm age, evaluated through two items adapted from Jaworski & Kohli (1993). The non-response bias for firm size and firm age was not significant.

### 3.2. Measures

Each item was measured using a five-point Likert-type scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The study examined existing literature to create items relevant to various aspects of intellectual capital, knowledge management capability, and innovation ambidexterity. Additionally, a structured survey tool was developed to collect the necessary data.

## 4. Results

### 4.1. Descriptive Statistics and Construct Validity

Prior to conducting data analysis, the Kaiser-Meyer-Olkin (KMO) test yielded a value of 0.694, indicating the suitability of the data for factor analysis of the variable constructs. The minimum threshold for

this test, as per Hair et al. (2006), is 0.50. Tables 1 and 2 present the means, standard deviations, and zero-correlation matrix for all variables.

Table 1 includes the mean, standard deviation, factor loadings and Cronbach's alphas. The average ratings for firms' intellectual capital, knowledge management capability, intangible resources advantage, and innovation ambidexterity were 4.015, 3.533, 3.993, and 4.062, respectively (with standard deviations of 0.441, 0.705, 0.457, and 0.444, respectively). These figures suggest a high level of competence in intellectual capital and knowledge management capability, indicating their significant role in achieving firm success. This analysis utilized factor loadings for each item greater than 1, employing a cutoff value of 0.40 as per Nunnally and Bernstein (1994). As summarized in Table 1, all reliability coefficients for the constructs (0.79–0.90) exceeded the 0.70 benchmark. The measured reliability for intellectual capital, knowledge management capability, innovation ambidexterity, and intangible resources advantage were 0.844, 0.901, 0.790, and 0.835, respectively.

Confirmatory factor analysis (CFA) was applied to evaluate convergent and discriminant validity. Convergent validity was assessed by examining whether the indicators of the same construct have high correlations among themselves. The loadings of each item should be higher than 0.7, indicating good convergent validity. For discriminant validity, all item loadings should have the strongest loadings on their corresponding construct compared with other constructs. The discriminant validity of all constructs was found to be good.

Table 2 presents the correlations between different variables. This study assessed construct reliability, as well as convergent and discriminant validity, to ensure that all reflective constructs are reliable and valid. To assess reliability and average



variance extracted (AVE) should be higher than 0.7 and 0.5, respectively. Table 2 shows that intellectual capital exhibits the strongest correlation with knowledge management capability ( $r = 0.352^{**}$ ,  $p \leq 0.001$ ). Additionally, knowledge management capability demonstrates the strongest correlation with innovation ambidexterity ( $r = 0.466^{**}$ ,  $p \leq 0.001$ ). After examining the relationships among each variable, the analysis reveals a spectrum of low to moderate, positive, and significant relationships. There is no evidence of multicollinearity among the variables, as indicated by variance inflation factor (VIF) values ranging from 1.004 to 1.525 (Hair et al., 2010). A VIF of 1.0 suggests the absence of multicollinearity, while a maximum VIF exceeding 10.0 indicates its presence. As depicted in Table 1, there are no issues of collinearity in this dataset, as the correlations between the independent variables are not high.

#### 4.2. Hypotheses Testing

Table 3 shows the hierarchical regression analysis between the independent and dependent variables. Both steps 2 and 4 provide strong support for the hypotheses. Specifically, the analysis revealed highly significant coefficients for the relationship between intellectual capital and knowledge management capability ( $\beta = 0.363$ ,  $p < 0.001$ ) and for

the relationship between knowledge management capability and innovation ambidexterity ( $\beta = 0.257$ ,  $p < 0.01$ ). This robust statistical evidence solidly supports hypotheses H1 and H2.

The analysis yielded an  $R^2$  value of 0.173, suggesting that 17.3% of the variance in knowledge management capability could be accounted for by intellectual capital. Additionally, 35.1% of the variation in knowledge management capability was found to be positively associated with innovation ambidexterity. Step 4 focused on testing the moderating influence of intangible resources advantage through interaction for hypothesis H3. The results from the moderated regression analysis in step 3 were statistically significant, leading to a significant increase in explained variance (adjusted  $R^2 = 0.348$ ). In this step, it was observed that intangible resources advantage had a positive moderating effect ( $\beta$  intangible resources advantage  $\times$  knowledge management capability = 0.128,  $p < 0.05$ ) on the interactive relationship between knowledge management capability and innovation ambidexterity.

Comparing these four steps revealed incremental increases in  $R^2$  at each stage of the hierarchical analysis, suggesting the direct effects of the independent variables (Cohen et al., 1983). To aid in interpreting the moderating effects, graphical representations of the interactions are presented.

**Table 1.** Summary statistics of the measurement analysis.

Variables	Mean	SD	Item loadings	Cronbach's alphas ( $\alpha$ )
Intangible resources advantage	3.993	0.457	0.818–0.863	0.790
Knowledge capability management	3.533	0.705	0.802–0.898	0.901
Intellectual capital	4.015	0.441	0.716–0.842	0.844
Innovation ambidexterity	4.062	0.444	0.734–0.849	0.835

Notes:  $^{**}$ Correlation is significant at the 0.01 level (2-tailed).

**Table 2.** Inter-factor correlations and related AVEs.

Construct	(1)	(2)	(3)	(4)
(1) Intangible resources advantage	0.879			
(2) Knowledge capability management	0.557 $^{**}$	0.927		
(3) Intellectual capital	0.765 $^{**}$	0.352 $^{**}$	0.893	
(4) Innovation ambidexterity	0.551 $^{**}$	0.466 $^{**}$	0.477 $^{**}$	0.888

Notes: The bold, underlined figures on the diagonal are AVEs.

**Table 3.** Hierarchical regression analysis.

	Knowledge capability management		Innovation ambidexterity		VIF
	Step 1	Step 2	Step 3	Step 4	
Firm size	-0.051	-0.064	0.057	0.051	1.199
Firm age	0.339	0.366	-0.134	-0.110	1.236
Intellectual capital		0.363***			1.004
Knowledge capability management			0.257**	0.302**	1.525
Intangible resources advantage			0.407***	0.383***	1.492
Knowledge capability management *Intangible resources advantage				0.128*	1.067
F	2.244	7.053**	13.545***	12.123***	
R <sup>2</sup>	0.042	0.173	0.351	0.380	
Adjusted R <sup>2</sup>	0.023	0.149	0.325	0.348	

Notes: \*\*\*Significant at the 0.001 level; \*\*Significant at the 0.01 level; \*Significant at the 0.05 level.

## 5. Discussion and Contributions

This study aimed to explore how intellectual capital and knowledge management capability influence each other, with a particular focus on the moderating role of intangible resources advantage in enhancing innovation ambidexterity. The results revealed distinct patterns in the adoption of knowledge management capability based on levels of intellectual capital (refer to Table 3). Notably, an intriguing finding was the positive and statistically significant interaction of intangible resources advantage on the relationship between knowledge management capability and innovation ambidexterity. To better illustrate this interaction, graphical representations of the relationship between knowledge management capability and innovation ambidexterity under varying levels of intangible resources advantage are provided in Table 3.

Following this, a detailed discussion on all three proposed hypotheses is presented. Hypothesis H1 suggested that intellectual capital significantly influences knowledge management capability. As demonstrated in regression step 2 of Table 3, this proposition received empirical support. Additionally, these results corroborated a prior study by Tsou and Chen (2020), which proposed that human capital (HC) positively impacts a firm's learning capabilities by enabling individuals to learn and apply their knowledge and experiences within organizations to generate innovative ideas. Indeed, strong human capital allows organizations to acquire new information and enhance individual abilities. This, in turn, creates opportunities for learning capability development, the integration of new knowledge with

existing knowledge, and reconfiguration in response to environmental changes (Altintas & Ambrosini, 2019).

Highly educated, skilled, and experienced workers are better able to recognize opportunities and risks and adapt to changing conditions. Employees' capacity to effectively integrate, reconfigure, and reallocate resources and capabilities, as well as learn, apply, and share essential and valuable information, are the primary drivers of this process.

The results of H2 indicated that knowledge management capability has a positive effect on innovation ambidexterity, as shown in regression step 3 of Table 3. Therefore, the research findings align with those of researchers like Soto-Acosta et al. (2018), who elucidated that organizations' capacity to employ exploratory and exploitative innovation is contingent upon their ability to promptly respond to external shifts brought about by changes in customer preferences, technological advancements, or fluctuations in product demand, in addition to the development of various internal capabilities like information technology and knowledge management. According to Del Giudice and Della Peruta (2016) and Abubakar et al. (2019), the proliferation of knowledge and knowledge management systems (KMSs) has led to the development of integrated and shared systems that enhance corporate performance.

Furthermore, while earlier research (López-Sáez et al., 2010) suggested that acquiring knowledge from external sources and assimilating and applying it aids companies in pursuing new market opportunities and improving innovation performance, the present study's findings provide a more nuanced understanding of

how knowledge management capabilities foster innovation ambidexterity. This result is consistent with the findings of innovation research, which suggests that knowledge is a company's most strategic source of capital for innovation potential (Martinez Conesa et al., 2017; Soto-Acosta et al., 2016).

To test the moderating role of intangible resources advantage in the relationship between knowledge management capability and innovation ambidexterity, the forecasting results indicate that the interaction effect of intangible resources advantage increases the predictive power from 32.5% to 34.8%. This demonstrates that the relationship between knowledge management capability and innovation ambidexterity significantly strengthens when an organization possesses an advantage in intangible resources, as shown in the analysis in step 4 of H3. Saunila and Ukko (2014) stated that innovation capabilities are largely derived from intangible assets, which Itami (1987) described as including employee know-how, managerial systems, company reputation, intellectual property, and informal social networks.

### 5.1. Practical Implications

The practical significance of this research lies in highlighting the vital connection between intellectual capital and knowledge management capability, which holds importance for both researchers and practitioners. The study underscores that intellectual capital plays a crucial role in knowledge management capability, and its integration with intangible resource advantage has a positive moderating effect. This finding offers valuable insights for firms aiming to enhance their understanding of the benefits associated with developing knowledge management capability and fostering innovation ambidexterity. Consequently, managers and organizations may face challenges in exerting managerial control to ensure the success of innovation ambidexterity.

The insights derived from the study of knowledge management capability, innovation ambidexterity, intellectual capital, and intangible competitive advantage offer several practical implications for organizations striving to maintain a competitive edge in today's dynamic business environment.

First, organizations should prioritize the development and enhancement of their intellectual capital. This includes continuous investment in employee training and development (human capital),

refining organizational processes and systems (structural capital), and fostering strong relationships with external stakeholders (relational capital). The Knowledge-based View (KBV) holds that knowledge management capability (KMC) is an essential component of organizations (Tseng, 2016) and plays a major role in the growth of innovation (Taghizadeh et al., 2020). By doing so, companies can create a robust foundation for effective knowledge management and innovation. Second, implementing comprehensive knowledge management systems is crucial. Organizations should develop systematic processes for creating, sharing, and utilizing knowledge. This can be achieved through the use of advanced technologies such as knowledge management software, intranets, and collaborative platforms. Encouraging a culture of knowledge sharing and continuous learning within the organization is also essential. Third, to achieve innovation ambidexterity, organizations need to balance exploratory and exploitative innovation activities. This requires a dual focus on exploring new opportunities and refining existing capabilities. Companies should establish dedicated teams or departments for both types of innovation, ensuring that resources and support are allocated appropriately to each. March (1991) asserts that innovation can produce desired performance benefits when exploration and exploitation are balanced. Fourth, organizations should recognize the value of intangible assets such as brand reputation, intellectual property, and customer relationships. These assets often provide a sustainable competitive advantage that is difficult for competitors to replicate. Effective management and strategic utilization of these intangible assets can enhance both exploratory and exploitative innovation efforts. Fifth, creating an organizational culture that supports innovation is vital. Leadership should encourage risk-taking, experimentation, and creativity. Providing employees with the freedom to explore new ideas and recognizing and rewarding innovative contributions can drive a culture of continuous improvement and innovation.

Therefore, organizations should establish metrics to assess the effectiveness of their knowledge management practices. This includes measuring the creation, sharing, and application of knowledge within the organization. Despite the fact that knowledge is a valuable resource with significant strategic potential for any business, that business needs a Knowledge Management Capability to assess and react swiftly to rivals' actions (Yang, 2020). Regular evaluation of

these metrics can help identify areas for improvement and ensure that knowledge management efforts are aligned with organizational goals. Besides that, integrating knowledge management and innovation strategies into the overall business strategy is essential. This ensures that knowledge management and innovation efforts are aligned with the organization's strategic objectives, enhancing coherence and effectiveness. Senior leadership should actively participate in and support these initiatives to demonstrate their importance.

## 5.2. Limitations and Future Research Directions

This study has certain limitations that open up avenues for future research. One limitation is the use of cross-sectional data, which prevents a clear determination of the temporal sequence in the relationships between intellectual capital, knowledge management capability, intangible resource advantage, and innovation ambidexterity. Therefore, developing a time-series analysis and conducting research within a longitudinal framework would offer greater insights into causality. Additionally, further exploration in different countries and contexts is necessary to generalize the findings and broaden the understanding of these relationships.

## Acknowledgments

The support by Faculty of Management Technology, Rajamangala University of Technology Isan, Surin Campus, Thailand is acknowledged.

## References

- Abubakar, A.M., Elrehail, H., Alatailat, M.A. & Elçi, A. (2019). Knowledge management, decisionmaking style and organizational performance, *Journal of Innovation and Knowledge*, 4(2), 104–114.
- Ahmed, S.S., Guozhu, J., Mubarik, S., Khan, M. & Khan, E. (2019). Intellectual capital and business performance: the role of dimensions of absorptive capacity, *Journal of Intellectual Capital*, 21(1), 23–39.
- Allameh, S.M. (2018). Antecedents and consequences of intellectual capital, *Journal of Intellectual Capital*, 19(5), 858–874.
- Altintas, G., & Ambrosini, V. (2019). *Dynamic Managerial Capabilities*, Oxford Research Encyclopedia of Business and Management.
- Armstrong, S. J. & Overton, T. S. (1977). Estimating non-response bias in mail survey, *Journal of Marketing Research*, 14, 396–420.
- Barney, J.B. (1991). Firm resources and sustained competitive advantage, *Journal of Management*, 17(1), 99–120.
- Carnabuci, G. & Operti, E. (2013). Where do firms' recombinant capabilities come from? Intraorganizational networks, knowledge, and firms' ability to innovate through technological recombination, *Strategic Management Journal*, 34(13), 1591–1613.
- Castaneda, D. I., & Cuellar, S. (2020). Knowledge sharing and innovation: A systematic organizational learning practices in innovation performance, *European Management Review*, 17(4), 835–855.
- Cohen, S., Kamarck, T. & Mermelstein, R. (1983). A global measure of perceived stress, *Journal of Health and Social Behavior*, 24, 385–396.
- Cuganesan, S. (2005). Intellectual capital-in-action and value creation, *Journal of Intellectual Capital*, 6(3), 357–373.
- Del Giudice, M. and Della Peruta, M.R. (2016). The impact of IT-based knowledge management systems on internal venturing and innovation: a structural equation modeling approach to corporate performance, *Journal of Knowledge Management*, 20(3), 484–498.
- Hair, J.F Jr., Black, W.C., Babin, B.J., Anderson, R.E. & Tathan, R.L. (2006). *Multivariate Data Analysis*, 6th ed. Upper Saddle River, New Jersey: Prentice-Hall.
- Hair, J.F Jr., Black, W.C., Babin, B.J. & Anderson, R.E. (2010). *Multivariate Data Analysis a Global Perspective*, Upper Saddle River, New Jersey: Prentice-Hall.

- Hall, R. (1992). The strategic analysis of intangible resources, *Strategic Management Journal*, 13(2): 135–144.
- Hofer, C. & Schendel, D. (1978). *Strategy formulation: Analysis concepts*, St. Paul, MN: West Publishing.
- Hsu, I. C., & Sabherwal, R. (2011). From intellectual capital to firm performance: The mediating role of knowledge management capabilities, *IEEE Transactions on Engineering Management*, 58(4), 626–634.
- Lei, H., Khamkhoutlavong, M. & Le, B.P. (2020). Fostering exploitative and exploratory innovation through HRM practices and knowledge management capability: the moderating effect of knowledge-centered culture, *Journal of Knowledge Management*, 25(8), 1926–1946.
- Jaworski, B. J., & Kohli, A. K. (1993). Market orientation: antecedents and consequences, *Journal of Marketing*, 57(3), 53–70.
- Itami, H. (1987). *Mobilizing invisible assets*, Harward University Press, Cambridge. MA.
- Kaufmann, L. & Schneider, Y. (2004). Intangibles: a synthesis of current research, *Journal of Intellectual Capital*, 5(3), 366–388.
- Kim, B.Y. & Oh, H. (2004). How do hotel firms obtain a competitive advantage?, *International Journal of Contemporary Hospitality Management*, 16(1), 65–71.
- Lee, J.Y. (2011), *Incremental Innovation and Radical Innovation: The Impacts of Human, Structural, Social, and Relational Capital Elements, Operations and Sourcing Management*. PhD Thesis, Michigan State University,
- López-Sáez, P., Navas-López, J., Martín-de-Castro, G., & Cruz-González, J. (2010). External knowledge acquisition processes in knowledge-intensive clusters, *Journal of Knowledge Management*, 14(5), 690–707.
- Lubatkin, M.H., Simsek, Z., Ling, Y. & Veiga, J.F. (2006). Ambidexterity and performance in smallto medium-sized firms: the pivotal role of top management team behavioral integration, *Journal of Management*, 32(5), 646–672.
- Machorro Ramos, F., Mercado Salgado, P., Cernas Ortiz, D.A. & Romero Ortiz, M.V. (2016). Influence of relational capital on the organizational performance of institutions of higher technological education, *Innovar*, 26(60), 35–50.
- Martinez-Conesa, I., Soto-Acosta, P. & Carayannis, E.G. (2017). On the path towards open innovation: assessing the role of knowledge management capability and environmental dynamism in SMEs, *Journal of Knowledge Management*, 21(3), 553–570.
- Melia, ' M. R., P'erez, A., B., & Dobon, ' S. R. (2010). The influence of innovation orientation on the internationalisation of SMEs in the service sector, *The Service Industries Journal*, 30(5), 777–791.
- Messeni Petruzzelli, A. (2019). Trading knowledge for status: Conceptualizing R&D alliance formation to achieve ambidexterity, *Technological Forecasting and Social Change*, 145, 36–42.
- March, J.G. (1991). Exploration and exploitation in organizational learning, *Organization Science*, 2(1), 71–87.
- Nunnally, J. C. (1978). *Psychometric Theory*, second ed. McGraw-Hill, New York.
- Nunnally, J.C. & Bernstein, I.H. (1994). *The Assessment of Reliability*, *Psychometric Theory*, 3, 248–292.
- Pedro, E., Leit~ao, J. & Alves, H. (2018). Intellectual capital and performance: taxonomy of components and multidimensional analysis axes, *Journal of Intellectual Capital*, 19(2), 407–452.
- Penrose, E.T. (1959). *The Theory of the growth of the firm*. Oxford University Press: New York.
- Saunila, M. & Ukko, J. (2014). Intangible aspect of innovative capability in SMEs: Impacts of size and industry, *Journal of Engineering and Technology Management*, 33, 32–46.

Soto-Acosta, P., Popa, S. & Palacios-Marques, D. (2016), E-business, organizational innovation and firm performance in manufacturing SMEs: an empirical study in Spain, *Technological and Economic Development of Economy*, 22(6), 885–904.

Soto-Acosta, P., Popa, S., & Martinez-Conesa, I. (2018). Information technology, knowledge management and environmental dynamism as drivers of innovation ambidexterity: A study in SMEs, *Journal of Knowledge Management*, 22(4), 824–849.

Taghizadeh, S.K., Karini, A., Nadarajah, G. & Nikbin, D. (2020). Knowledge management capability, environmental dynamism and innovation strategy in Malaysian firms, *Management Decision*, 59(6), 1386–1405.

Tseng, S.M. (2016). The effect of knowledge management capability and customer knowledge gaps on corporate performance, *Journal of Enterprise Information Management*, 29(1), 51–71.

Tsou, H. T., & Chen, J. S. (2020). Dynamic capabilities, human capital and service innovation: The case of Taiwan ICT industry,

*Asian Journal of Technology Innovation*, 28(2), 181–203.

Yang, X. (2020). Potential consequences of COVID-19 for sustainable meat consumption: the role of food safety concerns and responsibility attributions, *British Food Journal*, 123(2), 455–474.

Youndt, M.A., Subramaniam, M. & Snell, S.A. (2004). Intellectual capital profiles: an examination of investments and returns, *Journal of Management Studies*, 41(2), 335–361.

## AUTHOR BIOGRAPHIES



**Nirusa Sirivariskul** is an Assistant Professor in the Department of Management at the Faculty of Management Technology Rajamangala University of Technology Isan, Surin Campus, Thailand. With over 20 years of teaching experience, she specializes in subjects such as strategic management, management, change management, and other related fields. Her research interests encompass a wide range of topics, including innovation, entrepreneurship, change management, supply chain management, knowledge management, and intellectual capital.

# Custom hardware design for peripheral artery disease detection: Field-programmable gate arrays and application-specific integrated circuits

Nazarkar Pravalika<sup>1</sup>, A. Jabeena<sup>1\*</sup>, Vetriveeran Rajamani<sup>1</sup>

<sup>1</sup>School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

\* Corresponding author E-mail: ajabeena@vit.ac.in

(Received 2 September 2024; Final version received 21 October 2024; Accepted 30 October 2024)

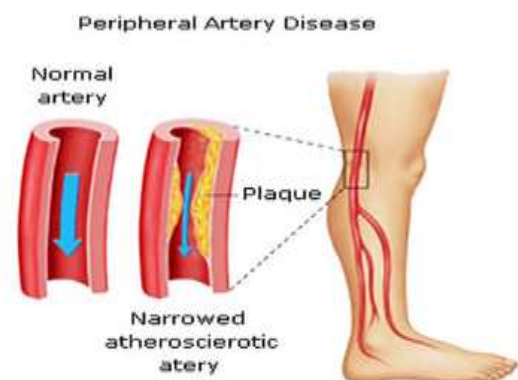
## Abstract

Atherosclerotic disorders, such as peripheral artery disease (PAD), have a significant negative impact on patient outcomes. Inadequate treatment and poor detection rates can result in cardiovascular complications and limb loss. There is great promise for improving the detection and treatment of PAD and other medical disorders through machine learning (ML) and artificial intelligence (AI) techniques. This paper highlights the use of field-programmable gate arrays (FPGAs) and application-specific integrated circuits (ASICs) to implement the fundamental ideas of AI and ML, specifically in the treatment of PAD. It emphasizes how these technologies can enhance drug selection, improve patient care, and refine disease phenotyping. This paper also describes how the integration of AI and ML with FPGA and ASIC technology can provide accurate and effective solutions to complex medical challenges, representing a significant breakthrough in medical analytics.

**Keywords:** Application-Specific Integrated Circuits, Field-Programmable Gate Arrays, Machine Learning, Peripheral Arterial Disease

## 1. Introduction

The cardiovascular condition known as peripheral artery disease (PAD) is caused by atherosclerosis, which restricts blood flow to the arteries and surrounding tissues (Campia et al., 2019; Criqui et al., 2015; Nordanstig et al., 2020; Rafnsson et al., 2020). After coronary artery disease and stroke, PAD is now the third most common atherosclerotic cardiovascular disease in terms of patient population (Criqui et al. 2015; Olin, 2000; Venkatesh et al., 2017). The most common symptom of PAD is intermittent claudication, an ischemic pain that arises when working leg muscles do not receive enough oxygen (Schorr et al., 2013). Over time, sedentary behavior tends to increase in PAD patients. Fig. 1. shows the difference between a normal artery and an artery with plaque.



**Fig. 1.** Normal artery and artery with plaque. (Reproduced from Dr. Abhilash, 15 December 2016)

Furthermore, functional impairment usually occurs before the diagnosis of PAD, and silent, undiagnosed PAD is linked to worse outcomes compared to intermittent claudication (Thrall et al.,



2018). It was estimated that more than 237 million adults aged 25 and older had PAD over the past decade, with prevalence sharply rising with age (Lareyre et al., 2023). Epidemiologic research also indicates a marked rise in the prevalence of PAD, particularly in low- and middle-income nations, suggesting the potential for a widespread PAD pandemic (Criqui & Aboyans, 2015). Treatments that are frequently employed include medication, surgery, and lifestyle changes; however, these may cause serious side effects and may not be suitable for everyone (Venkatesh et al., 2017). PAD risk is strongly associated with traditional cardiovascular risk factors, such as diabetes, smoking, and advanced age (Criqui & Aboyans 2015). Historically, PAD has received less attention than coronary artery disease and stroke, but in recent years, more focus has been paid to it, leading to new epidemiological advances (Olin 2000). A more severe form of cardiovascular disease that requires additional clinical care, is known as polyvascular disease. This disease is characterized by atherosclerosis in several artery beds. PAD can raise the risk of unfavorable outcomes by an equal amount or greater than that of heart disease or stroke (Schorr et al., 2013). The classification of PAD is shown in Fig. 2.

Early detection of PAD would allow for timely treatment that can slow the disease's progression, hence lowering the risk of major cardiovascular events. Nevertheless, in a primary care context, 40–60% of patients with PAD remain undiagnosed (Thrall et al., 2018). Ankle-brachial index (ABI) testing is the standard diagnostic procedure, but it is an expensive, highly specialized test that needs trained technologists in a vascular lab setting (Currie, 2019). Although physiological factors can impact the pulse wave recording technique, pulse wave measurements have shown potential for effectively detecting PAD, similar to ABI testing (Altman et al., 2024). Peripheral blood flow is necessary for a pulse wave to occur, and sympathetic nerve input, rather than vascular patency may influence pulse wave characteristics (Altman et al., 2024). Furthermore, by lowering blood flow, severe congestive heart failure can mimic inflow illness (Altman et al., 2024). Further investigation is required to establish the screening and diagnostic validity of pulse wave velocity measurements, which are a reliable hemodynamic measure for detecting PAD (Iglehart, 2006). Because PAD can mimic other conditions and is associated with aging, its diagnosis can be challenging. To help healthcare providers identify high-risk patients in their everyday clinical

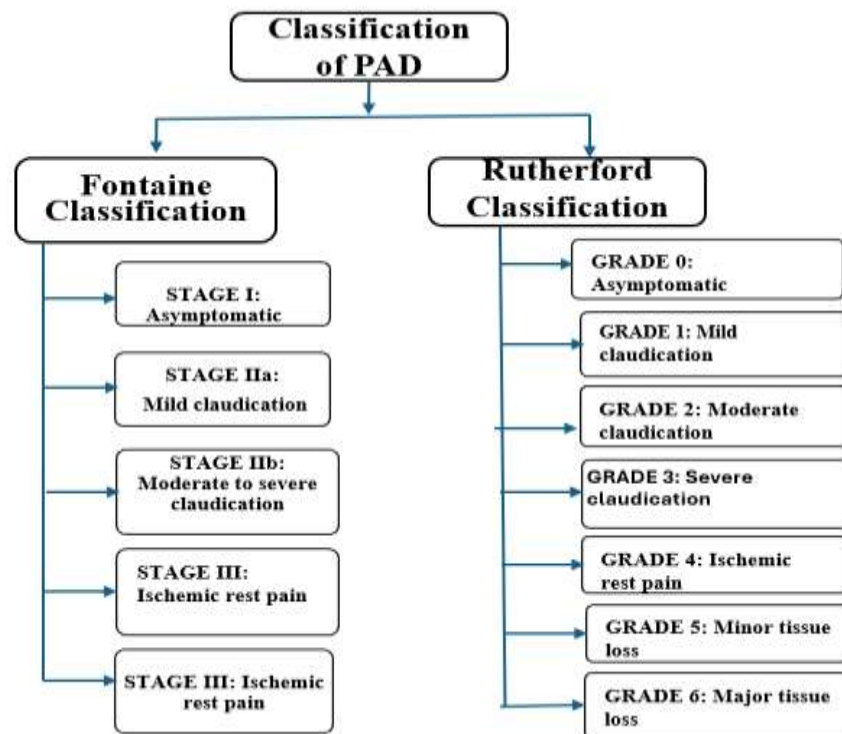


Fig. 2. Classification of peripheral arterial diseases (PAD).

practice, a non-invasive screening method is needed. The symptoms of PAD include claudication, unusual leg discomfort, and critical limb ischemia, which can result in ulcerations and possibly amputation. PAD is defined by an ABI of less than or equal to 0.90 (Nordanstig et al., 2020).

### 1.1. Prevalence and Clinical Significance of Peripheral Artery Disease

*Prevalence:* PAD involves the atherosclerotic narrowing or occlusion of arteries other than the coronary arteries and the aorta. It affects various arteries, including the carotid, vertebral, mesenteric, renal, and limb arteries.

*Risk Factors:* Common risk factors include smoking, hypertension, dyslipidemia, diabetes, and autoimmune/inflammatory conditions.

*Clinical significance:* PAD is associated with increased cardiovascular morbidity and mortality, and it significantly impairs quality of life.

*Screening and diagnosis:* Screening for PAD involves assessing risk factors and symptoms, as well as using tools like the ABI and various imaging modalities.

### 1.2. Challenges in Current Diagnostic Methods

Although PAD, often caused by atherosclerosis, can affect any artery outside the heart and brain, it is frequently linked to conditions of the lower limbs. The prevalence of PAD is increasing worldwide, especially in low- and middle-income nations, with significant rises in cases (Aboyans et al., 2017). Among the risk factors include diabetes, hypertension, and smoking (Criqui & Aboyans 2015). Chronic limb-threatening ischemia and intermittent claudication are two possible indications of PAD (Aboyans et al., 2018). Management of PADs typically involve secondary prevention, risk factor elimination, and early diagnosis (Venkatesh et al., 2017). In severe cases, revascularization may be necessary in addition to medications and lifestyle modifications (Adegbola et al., 2022).

### 1.3. Invasive and Non-Invasive Methods for Peripheral Artery Disease Diagnosis

Since the 1670s, various invasive and non-invasive diagnostic tools have been developed to address peripheral vascular diseases. These tools help in precise diagnosis and treatment planning. Imaging

methods for PAD are categorized into non-invasive and invasive modalities, each with distinct advantages, disadvantages, and associated costs (Lareyre et al., 2023).

#### 1.3.1. Non-Invasive Methods

##### 1. Color duplex ultrasound

Benefits: Non-invasive, non-ionizing, requires no contrast agent, provides hemodynamic data, and is useful for follow-up.

Cons: Requires an operator, it has a narrow range of vision; takes longer to examine; and can't evaluate calcified vessels as thoroughly.

##### 2. Computed tomography (CT) angiography

Benefits: Non-invasive, low cost, high availability, quick three-dimensional (3D) imaging, excellent spatial resolution, and the ability to post-process data.

Cons: Limited evaluation of calcified and infrapopliteal vessels; iodinated contrast material is necessary; ionizing radiation is used.

##### 3. CT with dual energy

Benefits: High spatial resolution, exceptional image quality in remote areas, and the ability to examine tissue in greater detail.

Cons: Limited accessibility and requires specialist equipment; ionizing radiation is used.

##### 4. CT using photon counting

Benefits: Minimizes blooming artifacts, provides intrinsic spectral information, results in minimal radiation exposure, offers excellent contrast-to-noise ratio, and has high spatial resolution.

Cons: Restricted accessibility and involves ionizing radiation.

##### 5. Magnetic resonance imaging (MRI)

Benefits: Non-invasive, does not require contrast material, offers high resolution, provides flow-independent assessment of vessels below the knee, delivers excellent soft tissue contrast for evaluating plaque, provide hemodynamic information, and uses gadolinium-based contrast that is more tolerable for patients with impaired renal function.

Cons: Greater cost compared to CT; longer acquisition time; limited assessment of calcifications due to certain procedures; claustrophobic difficulties; limitations with non-MRI conditional devices.

#### 1.3.2. Invasive Methods

##### 1. Digital subtraction angiography

Advantages: High resolution, fast.

Disadvantages: Invasive, requires iodinated contrast material, limited ability to assess vessel wall.

#### 2. Intravascular ultrasound

Advantages: Provides widespread, detailed diagnostic information on lumen size, vessel wall, and plaque burden.

Disadvantages: Susceptible to artifacts, lower frame rate, operator dependent.

#### 3. Optical coherence tomography

Advantages: High resolution, provides both two-dimensional and 3D images, suitable for smaller vessels.

Disadvantages: Limited penetrative depth, restricted field of view, requires saline irrigation accompanied by inflow occlusion.

#### 4. Angioscopy

Advantages: Direct visualization of vessel wall and wall-associated structures, provides colored images.

Disadvantages: Cannot measure disease presence, plaque volume, content, or depth; requires saline irrigation accompanied by occlusion.

This comprehensive comparison aids in selecting the most appropriate imaging modality based on clinical needs, patient condition, and available resources (Lareyre et al., 2023).

### 1.4. Advantages over Traditional Methods

the Management of PAD involves revascularization, advanced diagnostics, and post-care to prevent complications. Challenges include embolization, calcification, and restenosis. New treatments feature less invasive methods, drug-eluting technologies, and biomimetic stents, enhancing outcomes. Machine learning (ML) and real-time data improve early diagnosis and treatment of PAD (Beckman et al., 2021). Nanotechnology in testing offers cost-effective, faster, more accurate, and sensitive solutions, aiding in the identification and management of PAD (Geiss et al., 2019). With these technological advances, PAD can now be diagnosed and treated with greater accuracy, potentially improving patient outcomes and quality of life (Elbadawi et al., 2021).

Building on the progress in PAD treatment, Section 2 provides an overview of ML, including its various types. In Section 3, we explore how ML is applied in the medical field, particularly for diagnosing PAD. Section 4 focuses on the use of ML with field-programmable gate arrays (FPGA) technology for PAD diagnosis, discussing both its advancements and limitations. Section 5 examines

how ML integrates with artificial intelligence (AI) and application-specific integrated circuits (ASIC) technologies in medical diagnosis, highlighting improvements with AI performance and efficiency for PAD diagnosis. The review will conclude with proposed work and final thoughts.

## 2. Artificial Intelligence Technologies in Healthcare

Artificial intelligence and ML can greatly enhance healthcare by predicting illness outcomes, patient readmissions, and therapy responses (Jiang et al., 2017). Deep learning improves diagnostic accuracy in medical image analysis using MRIs, CT scans, and X-rays (Litjens et al., 2017). Natural language processing extracts valuable information from clinical documentation, improving decision-making and data accuracy. AI in genomic medicine identifies genetic markers, understands genetic variations, and develops personalized therapies, thus advancing precision medicine. ML accelerates drug discovery, reducing risk and improving decision-making in target validation and drug design. In personalized medicine, AI tailors treatments based on genetic profiles and previous responses, optimizing success rates and minimizing adverse effects (Kourou et al., 2015).

Artificial intelligence in robotic surgery enhances precision, reduces errors, and improves patient outcomes (Hashimoto et al., 2018). Wearable sensors enable remote patient monitoring, with AI analyzing data and alerting healthcare providers to health trends (Rojas & Wang, 2020). AI also improves electronic health record (EHR) management, enhancing data accuracy, care coordination, and administrative efficiency. Addressing concerns about patient privacy, bias, and transparency is crucial for the ethical use of AI in healthcare (Rojas & Wang, 2020).

### 2.1. Machine Learning and Its Applications in Medical Diagnosis of Peripheral Artery Disease

The ability of machine to independently simulate intelligent activity using ML has significantly advanced in the field of computer science (Bini, 2018), leading to its increasing application in various fields, including medicine. The use of AI in medicine has expanded rapidly, with collaborations between the medical field and AI garnering significant attention from the global economy, particularly in 2016 (Buch et al., 2018). AI's role in medicine primarily involves

automating diagnostic procedures and managing patient care, which allows medical professionals to focus on more complex, non-automatable tasks.

In medical applications, ML is typically categorized as supervised learning, where output variables are predicted from input data, and unsupervised learning, which involves clustering different groups based on specific interventions. This growing utilization of ML is not limited to direct medical care but extends to areas like human resources, allowing practitioners and specialists to build sophisticated models and extract valuable medical knowledge.

Machine learning predictive models are particularly useful in clinical settings, where they can improve decision-making and even autonomously diagnose various disorders (Criqui et al., 2015; Schorr et al., 2013). Furthermore, corporations leveraging ML in drug prescription can help physicians identify new medical opportunities, potentially saving lives by accurately detecting pathologies (Lo et al., 2018). Additionally, by reducing medical expenses, stabilizing patient flow, and improving data quality, ML models offer a more effective alternative to traditional diagnostic methods (Napolitano et al., 2016).

The next section explores how ML is applied to PAD, how it enhances diagnostic accuracy, facilitates early detection, and ultimately improves patient outcomes. By analyzing large sets of vascular data, ML models are instrumental in identifying patterns that may go unnoticed using traditional diagnostic methods. This section delves into the effectiveness of these techniques, providing actionable insights that support clinicians in making more informed decisions regarding PAD treatment and management.

It also highlights how ML could enhance clinical care and medical research, especially when applied to electronic health data (Jordan & Mitchell, 2015). Moreover, ML methods are emphasized for their capacity to identify illnesses and forecast health outcomes. For example, predicting the course of diabetes from electronic health information and classifying skin cancer from photos are two notable applications (Esteva et al. 2017).

Researchers have gained new insights from clinical incident reports by integrating ML with natural language processing strategies (Ong et al., 2012) integrated in social networking sites. Evaluations of physician performance and patient testimonials following beneficial cancer therapies have been enhanced using ML (Ong et al., 2012).

Current PAD treatments often use a generic approach, but research is exploring more options beyond aspirin, statins, and smoking cessation, despite their higher costs and risks (Flores et al., 2021). ML models can optimize PAD treatment for patients with comorbidities by analyzing drug interactions and polypharmacy side effects. Training these algorithms on PAD patients data can help create synchronized treatment plans, replacing reductionist approaches with AI that identifies PAD subgroups and integrates polypharmacy and pharmacogenetics data. Advanced data science can also assess long-term therapy safety in real-world contexts, addressing clinical trial exclusions, like those with congestive heart failure on cilostazol, which showed no adverse effects (Brass et al., 2006).

Machine learning can identify complex PAD risk factors using EHR data but requires structured data and portable analysis pipelines. Improved prediction models, such as those for surgical site infections and limb ischemia post-revascularization, can enhance PAD treatment and outcomes (Brass et al., 2006). Combining clinical and imaging data, like Doppler waveforms and CT angiograms, using ML and computer vision can further improve PAD diagnosis and reduce invasive procedures (Misra et al. 2019). Different researchers have applied various ML algorithms to a range of diseases and have assessed their respective advantages and disadvantages. For diseases similar to PAD, these findings have been summarized in Table 1, highlighting the advantages and disadvantages of each algorithm.

Fig. 3 depicts the number of papers published on ML for PAD from 2002 to 2024, with a substantial upward trend. The data, derived from the Scopus database, reveal consistent growth in publications over time, with a notable surge in recent years as the use of ML in healthcare has gained traction.

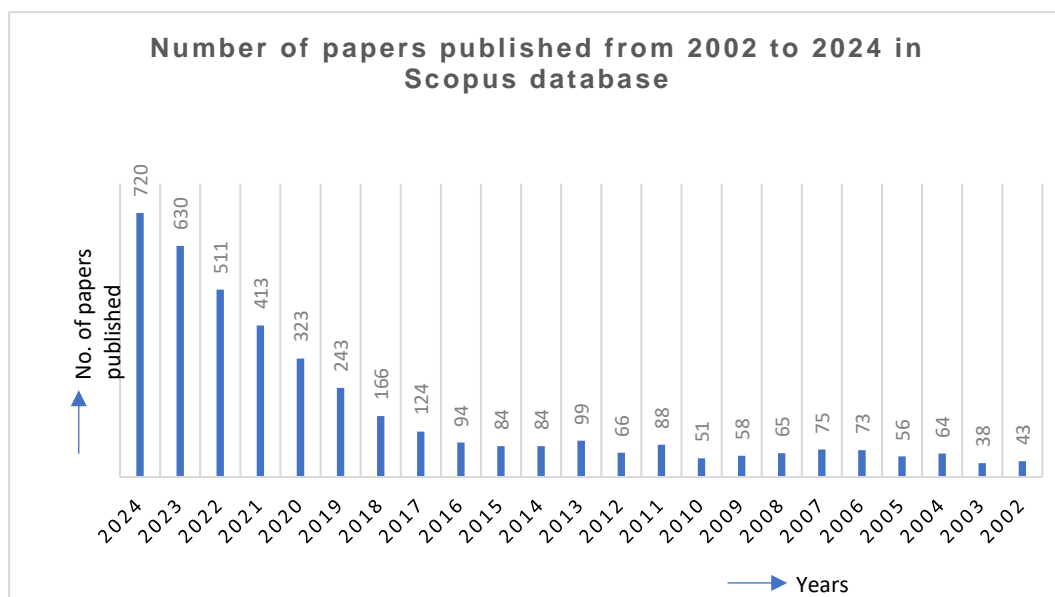
## 2.2. Types of Machine Learning Techniques Used

Machine learning techniques are becoming increasingly effective in risk assessment, disease prognosis, and image-based diagnosis. With a wide range of applications, ML is one of the fastest-growing fields in computer science. It involves the automatic identification of significant patterns in data. Giving algorithms the capacity to learn and adapt is the focus of ML tools (Shalev-Shwartz & Ben-David 2014). ML algorithms are classified based on their intended outcomes. Supervised learning converts inputs into desired outputs and is common due to its role in

**Table 1.** Types of machine learning algorithms and their advantages and disadvantages.

ML algo-rithm	Advantages	Disadvantages
KNN	<ul style="list-style-type: none"> <li>- A model that is inexpensive and simple to use.</li> <li>- Is utilized in both regression and classification.</li> <li>- Handles multiclass situations seamlessly.</li> </ul>	<ul style="list-style-type: none"> <li>- The computation is quite high.</li> <li>- Classification costs for unknown records are comparatively high.</li> <li>- Elevated sensitivity to irrelevant features.</li> </ul>
K-Means	<ul style="list-style-type: none"> <li>- Simple to execute.</li> <li>- More effective when variables are larger than hierarchical clustering.</li> </ul>	<ul style="list-style-type: none"> <li>- It is challenging to estimate the K value.</li> <li>- Performance declines with a globular cluster.</li> <li>- Sensitive to noise and anomalies.</li> </ul>
SVM	<ul style="list-style-type: none"> <li>- Can handle linear and nonlinear data.</li> <li>- Lower likelihood of overfitting.</li> <li>- Scales well with high-dimensional data.</li> </ul>	<ul style="list-style-type: none"> <li>- Performance degrades when dealing with huge datasets.</li> <li>- Choosing a good kernel function is challenging.</li> <li>- Less effective in noisy datasets.</li> </ul>
Naïve Bayes	<ul style="list-style-type: none"> <li>- Convenient for large datasets.</li> <li>- Handles both discrete and continuous data.</li> <li>- Can be applied to both multiple and binary classification.</li> <li>- Insensitive to irrelevant features.</li> </ul>	<ul style="list-style-type: none"> <li>- Computationally demanding, particularly for models with many variables.</li> <li>- Models that have been properly trained and applied may occasionally underperform.</li> <li>- Lack complexity.</li> </ul>
Logistic regression	<ul style="list-style-type: none"> <li>- Computationally effective.</li> <li>- Simple regularization.</li> <li>- No scaling is needed for input features.</li> </ul>	<ul style="list-style-type: none"> <li>- Solving a nonlinear problem is challenging.</li> <li>- Risk of overfitting.</li> </ul>
Decision tree	<ul style="list-style-type: none"> <li>- Utilized for classification as well as regression.</li> <li>- Simple management of both categorical and numerical data.</li> </ul>	<ul style="list-style-type: none"> <li>- Overfitting could happen if the tree is constructed repeatedly.</li> <li>- Larger trees become challenging to understand.</li> </ul>
Random Forest	<ul style="list-style-type: none"> <li>- Applicable to both classification and regression problems.</li> <li>- Solves overfitting issues in a decision tree.</li> </ul>	<ul style="list-style-type: none"> <li>- Training takes a long time.</li> <li>- Complexity increases.</li> </ul>
Deep learning	<ul style="list-style-type: none"> <li>- Automatically identifies features.</li> <li>- Applicable to several types of data.</li> </ul>	<ul style="list-style-type: none"> <li>- For training, GPUs are required.</li> <li>- Complicated data models make training extremely expensive.</li> </ul>

Abbreviations: GPU: Graphics Processing Units; KNN: K-Nearest Neighbors; SVM: Support Vector Machine.



**Fig. 3.** Number of papers published on machine learning for peripheral arterial disease from 2002 to 2024. Source: <http://www.webofscience.com>

teaching classification systems. The complexity of data generation has led to the adoption of both supervised and unsupervised techniques (Shalev-Shwartz & Ben-David, 2014), enabling accurate predictions on unseen data (Sarker, 2021).

### 2.2.1. Unsupervised Learning and Semi-Supervised Learning

Unsupervised learning allows machines to recognize patterns in unlabeled data, making it useful for tasks like clustering. Semi-supervised learning combines labelled and unlabeled data. In this approach, a supervised algorithm is trained on labeled data, while an unsupervised algorithm is used to label new instances for further training (Smiti, 2020).

### 2.2.2. Reinforcement Learning

By using prior experience, this ML technique enables robots or agents to determine the most optimal behavior in a given scenario. Machines gather information during training to enhance their functionality (Smiti, 2020).

Each ML algorithm has a unique set of benefits and limitations. While some are more focused on specificity or computing efficiency, others are especially excellent in terms of accuracy and sensitivity. The trade-offs highlighted in these comparisons underscore how different algorithms perform better for different datasets or tasks. Therefore, the ideal approach to use relies on the unique requirements and limitations of the application. To analyse the efficacy of ML algorithms, a number of researchers have investigated their application across

a variety of datasets and assessed important performance measures like accuracy, sensitivity, specificity, and F1 score. Both Random Forest and Support Vector Machine (SVM) algorithms are ideal due to their high accuracy and ability to handle complex data, with Random Forest effectively mitigating overfitting and SVM adept at managing high-dimensional data.

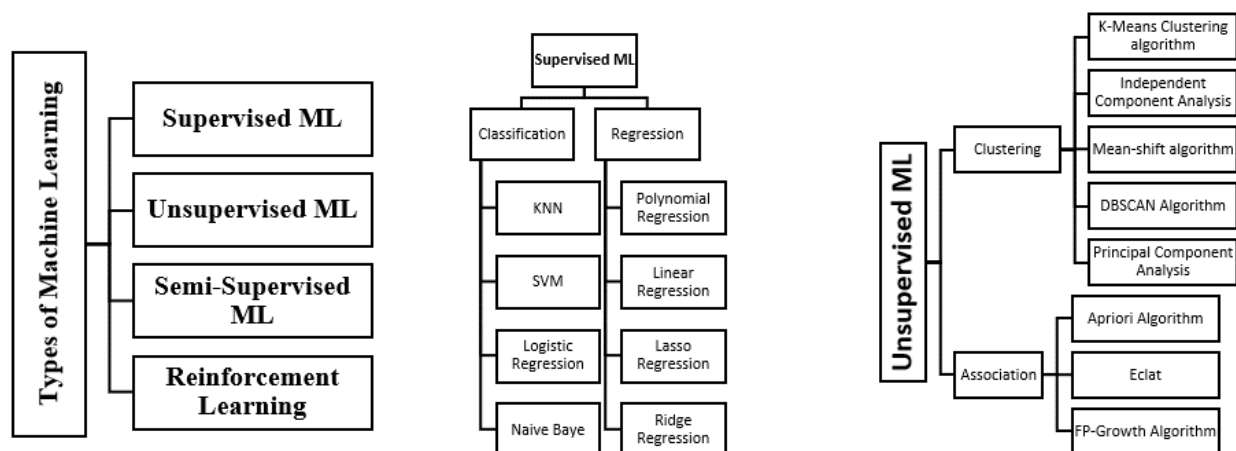
### 2.3. Future-Challenges and Concepts to Consider

*Prospective assessment:* ML and AI in vascular medicine show great promise but need real-world validation and comparison with standard care to avoid overdiagnosis and overtreatment. Improved diagnostic accuracy from EHRs could enhance treatment adherence (Rolls et al., 2016)

*Data interoperability:* Large, diverse datasets are essential for ML and AI development. Challenges in data interoperability among EHRs, hospitals, and health systems can be mitigated by Common Data Models and Fast Healthcare Interoperability Resources. Wider acceptance of these standards is needed (Rolls et al., 2016).

*Algorithm inaccuracy:* Addressing programmed bias is crucial, as ML algorithms can perpetuate existing biases, especially affecting marginalized groups like Black and Native American people. Diverse datasets and transparency are essential to prevent this risk (Rolls et al., 2016).

*Confidentiality and safety:* Patient privacy must be protected. Blockchain and federated learning are being explored to securely combine data and build models while maintaining privacy (Yang et al., 2019).



**Fig. 4.** Classification of machine learning (ML). Abbreviations: DBSCAN: Density-Based Spatial Clustering of Applications with Noise; FP: Frequent Pattern; KNN: K-Nearest Neighbors; SVM: Support Vector Machine.



**Execution:** Early collaboration among data scientists, clinicians, and implementation scientists are key for intuitive AI tool integration. Design thinking, coupled with consideration of end-user needs at all stages, can prevent workflow issues like EHR alert fatigue (Li et al., 2020).

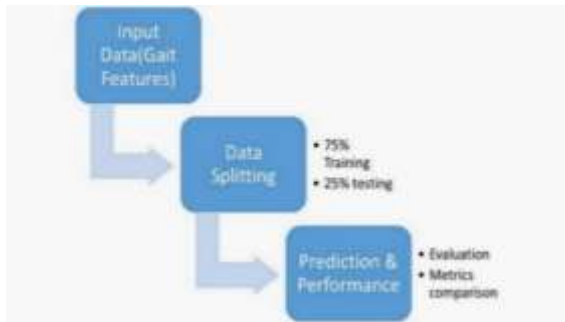


Fig. 5. Machine learning process.

### 3. Field-Programmable Gate Arrays in Peripheral Artery Disease Diagnosis

Among the more contemporary methods, the FPGA-based hardware solution has emerged as one of

the most commonly utilized in the healthcare field. However, the restructuring of material and temporal limits of current technologies presents a significant challenge to these methodologies. Despite this, FPGA circuits are favored due to their parallel and connected designs, low cost, reconfigurable features, and special-purpose circuits. In particular, a real-time architecture built on FPGA is proposed to speed up feature extraction as well as initial diagnosis (Gu et al., 2016).

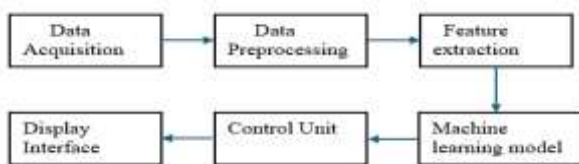
Application-specific integrated circuits, on the other hand, provide excellent performance and low power consumption, as they are designed for specific uses. ASICs are particularly essential for integrating unique designs into medical imaging equipment, ultimately leading to smaller form factors and lower costs for designs produced in large quantities (Talib et al., 2021). Similarly, FPGAs are highly useful in the medical field due to their post-fabrication reconfiguration capability. Owing to their adaptability to design changes and enhancements, FPGAs are often recommended for systems that require high performance, particularly in cases with real-time processing demands (Talib et al., 2021).

Table 2. Comparison of the results based on various machine learning techniques.

Authors	Machine learning techniques used	Number of datasets used	Accuracy	Sensitivity	Specificity	F1 score
Weissler et al., 2023	Not specified	Large cohort dataset	86%	87%	85%	0.84
Tomkins et al., 2023	Random Forest	Single large cohort	88%	81%	90%	0.76
Hogg et al., 2023	Not specified	2/3 training, 1/3 test set	85%	83%	87%	0.8
Jana et al., 2022	SVM, KNN	125 spectrograms	90.40%	79.97%	88.5–99%	-
Forghani et al., 2021	Genetic algorithm, RU-Boost	14 PAD patients, 19 healthy individuals	91.40%	90.00%	92.10%	-
Friberg et al., 2022	NLP, Random forest	800 ABI values	89%	-	-	-
Huthart et al., 2022	DUS	250 patient data	86%	81.00%	86.30%	-
Forghani et al., 2021	Deep Learning: BiLSM	14 PAD patients, 14 healthy individuals	94.80%	90.00%	97.40%	0.89
Zhang et al., 2022	LR, RF, XGBoost, LightGBM	Not specified	-	77%	72%-75%	-
Mistelbauer et al., 2021	CNN, RNN, U-Net, Deep vessel net	7,000 datasets	99.90%	92.90%	-	-
Gao et al., 2022	LR, RF	539 patients	-	100%	90.30%	-
Sasikala & Mohana-rathinam, 2024	KNN, LR, SVM, EDT, SGD, XG-Boost (HB+SMOTE+ED T)	Cleveland dataset	99.20%	98.70%	99.12%	99
		Statlog dataset	98.52%	98.13%	98.72%	98.09

Abbreviations: ABI: Ankle-brachial index; BiLSM: Bidirectional Long Short-Term Memory; CNN: Convolutional Neural Network; EDT: Elastic Distributed Training; KNN: K-Nearest Neighbors; LR: Logistic regression; NLP: Natural language processing; PAD: Peripheral artery disease; RF: Random Forest; RNN: Recurrent Neural Network; SGD: Stochastic Gradient Descent; SMOTE: Synthetic Minority Oversampling Technique; SVM: Support Vector Machine.

Moreover, FPGA technology is rapidly advancing in the areas of medical imaging and signal processing, significantly enhancing both accuracy and reliability. With their ability to process large volumes of data quickly and implement real-time image processing algorithms, FPGAs play a crucial role in aiding doctors with patient diagnoses. Additionally, these advancements are improving signal processing techniques, such as digital filtering and image recognition, making FPGAs powerful tools for the future of patient care (Ramya, 2024). Thus, medical imaging increasingly relies on FPGAs to enhance performance and versatility. Through their capabilities, FPGAs improve processing speed, accuracy, and resilience, offering customized, cost-effective solutions. They are especially effective in handling data-intensive tasks, reconfigurability, high performance, low power consumption, and cost-saving features, all of which enable real-time, accurate signal processing (Ramya, 2024).



**Fig. 6.** Block diagram for implementing medical image processing with machine learning on a field-programmable gate array.

Medical image processing with ML on an FPGA involves obtaining a raw medical image dataset, preprocessing it to improve quality using noise reduction and normalization, extracting features like edges and textures using edge detection and Histogram of Oriented Gradient algorithms, classifying or detecting results using trained convolutional neural networks (CNNs) or SVMs, and displaying the output on a user interface.

### 3.1. Advance Techniques Using Field-Programmable Gate Arrays

Field-programmable gate array capabilities capture the impact of enhanced control techniques, real-time simulation, and electronic instrumentation in areas like power systems, robotics, and mechatronics. Additionally, features such as hard memory controllers, analog resources, and floating-point operators offer significant benefits to designers, enabling more efficient and effective system designs (Rodríguez-

Andina *et al.*, 2015). The integration of soft processor cores, embedded processors, and high-performance hardware peripherals within FPGAs facilitates the creation of powerful system-on-chip platforms. This trend is exemplified by the rise of FPGA-based systems on chip, which feature optimized architectures that enhance connectivity across various applications (Rodríguez-Andina *et al.*, 2015). Moreover, new digital signal processing blocks, such as Altera's variable precision digital signal processing blocks, specifically address fixed-point challenges, alleviating resource consumption and latency issues caused by mantissa alignment and normalization in traditional designs (Altera Corporation 2015).

For instance, Ahmed *et al.* (2017) developed an FPGA-based system for the real-time detection of PAD using Doppler ultrasonography signals. Tested on 150 patients (75 with PAD and 75 healthy), the system achieved 92% sensitivity and 89% specificity, further underscoring the FPGA's real-time processing capability for clinical use.

Smith *et al.* (2019) implemented an ML technique for PAD detection on an FPGA platform using an SVM classifier trained on Doppler ultrasonography signals. They tested it on 200 patients (100 with PAD and 100 healthy), achieving 94% accuracy. The study highlighted the FPGA's fast processing and high accuracy, making it suitable for point-of-care diagnostics.

Lee *et al.* (2021) developed a high-speed FPGA-based system for diagnosing PAD using Doppler spectrograms. The system employed peak detection and fast Fourier transform for spectral analysis. Tested on 180 patients (90 with PAD and 90 healthy), it achieved 91% diagnostic accuracy, demonstrating faster processing than traditional methods and highlighting FPGA's potential for efficient PAD detection in clinical settings.

Zhang *et al.* (2022) developed a portable FPGA-based device for PAD evaluation, integrating Doppler ultrasound signal acquisition, processing, and display. Tested on 120 patients (60 with PAD and 60 healthy), it achieved 88% sensitivity and 85% specificity. The study highlighted the potential of portable FPGA devices for PAD screening, especially in remote or resource-limited areas.

Patel *et al.* (2016) developed an FPGA-based adaptive filtering system to enhance Doppler ultrasound signals for PAD detection. Using adaptive noise cancellation, the system improved signal clarity. Tested on 100 patients (70 with PAD and 30 healthy), it achieved 90% diagnostic accuracy, demonstrating

FPGA's effectiveness in real-time signal enhancement for accurate PAD diagnosis.

Gupta *et al.* (2018) developed an FPGA-based system for real-time hemodynamic analysis to detect PAD. Using Doppler ultrasound signals from 160 patients (80 with PAD and 80 healthy), the system achieved 89% sensitivity and 87% specificity. The study highlighted FPGA's capability for quick and accurate PAD diagnosis.

Kim *et al.* (2020) developed an FPGA-based system using a CNN for PAD detection. Tested on 190 patients (95 with PAD and 95 healthy), it achieved 95% accuracy. The study demonstrated the integration of deep learning models with FPGA for high-accuracy, real-time PAD diagnosis.

Zhao *et al.* (2021) developed a low-power FPGA-based system for continuous PAD monitoring. Tested on 130 patients (65 with PAD and 65 healthy), it achieved 88% diagnostic accuracy. The study highlighted the importance of power efficiency in wearable and portable devices for PAD monitoring.

Thompson *et al.* (2022) developed an FPGA-based system integrating Doppler ultrasound and photoplethysmography for comprehensive PAD diagnosis. Tested on 150 patients (75 with PAD and 75 healthy), it achieved 91% sensitivity and 90% specificity. The multi-sensor approach improved diagnostic accuracy and provided a thorough evaluation of PAD.

PAD can also be analyzed using gait features. One study, in particular, utilized gait features to identify different stages of PAD. Fig. 5 illustrates the ML approach employed for PAD detection using these gait features.

### 3.2. Challenges of Using Field-Programmable Gate Arrays in Medical Diagnosis

*Complexity:* FPGAs require deep understanding of Hardware Description Languages like VHDL and Verilog, deterring those accustomed to software solutions (Xie *et al.*, 2019).

*Power consumption:* High power usage makes FPGAs less suitable for portable devices (Lee *et al.*, 2020).

*Processing speed:* FPGAs are slower than ASICs, which can be a limitation for real-time applications (Banerjee *et al.*, 2018).

*High start-up costs:* Significant initial investments are needed for design tools, intellectual property cores, and skilled personnel (Huang *et al.*, 2021).

*Scalability issues:* Limited on-chip resources hinder FPGA's ability in handling larger datasets or complex algorithms (Chen *et al.*, 2017).

*Maintenance and upgrades:* More effort is required compared to software-based systems (Patel *et al.*, 2022).

While FPGAs face challenges such as complexity, high power consumption, and scalability issues that can hinder their effectiveness in medical diagnosis, ASICs present a compelling alternative by offering tailored performance, lower power usage, and enhanced efficiency for specific applications.

### 4. Application-Specific Integrated Circuits in Medical Imaging

Application-specific integrated circuits are uniquely designed for specific uses, offering excellent performance, smaller form factors, and lower power consumption (Munn *et al.*, 2011). They are crucial for medical imaging, enhancing computation and enabling parallel tasks (Alcaín, *et al.*, 2021). ASICs reduce costs for high-volume designs and extend equipment life, making healthcare more accessible (Alcaín *et al.*, 2017). Modern ASIC architectures provide competitive image processing with high-speed input/output, dedicated memory, and greater logic density, enabling novel medical imaging applications (Beyer *et al.*, 2009).

Artificial intelligence technologies like deep learning, ML, and neural networks are revolutionizing medical imaging. They improve visual recognition and data insights, enhancing efficiency, quality, and outcomes. Artificial neural networks use layers of nodes to process images, while deep learning, a more advanced form, uses multiple layers for detailed analysis (Langlotz *et al.*, 2019). As a subset of artificial neural networks, CNNs directly extract features for classification from images (Liew, 2018). Radiomics is the extraction of outcome-related imaging features to improve precision medicine (Thrall *et al.*, 2018). AI can identify key features and their combinations for predictive power, reducing redundancy in mathematical modeling. When integrating AI in medical imaging, it's crucial to consider regulations and ethics. Prioritizing patient-centered design ensures ethical and sustainable AI use in healthcare (Currie, 2019).

Application-specific integrated circuits and FPGAs each have benefits and limitations, and the best choice depends on specific medical imaging requirements and system specifications. The need for

**Table 3.** Main features of field-programmable gate array (FPGA) applications in medicine, with an emphasis on peripheral artery disease (PAD) diagnostics and other healthcare issues.

Paper	FPGA implementation	Application area	Performance	Challenges/Considerations
Ramya, 2024	FPGA in medical imaging	Medical imaging and signal processing	Improved accuracy and speed	Design complexity in medical applications
Samanta et al., 2023	Memristor-based logic gates	Hybrid logic gates for AI	Energy-efficient logic design	Integrating memristor technology into VLSI
Nagarajan et al., 2011	Pattern-based decomposition	Machine learning algorithms	Accelerated ML processing	Efficient pattern matching and decomposition
Rodríguez-Andina et al., 2015	Advanced FPGA features	Industrial applications of FPGAs	High flexibility and performance	Application-specific FPGA tuning
Altera Corporation, 2015	Arria 10 FPGA	General-purpose I/O for various tasks	High throughput	Optimizing power consumption and performance balance
Ahmed et al., 2017	Real-time detection of PAD	PAD diagnosis using FPGAs	Real-time, accurate PAD detection	Optimizing system for real-time use
Smith et al., 2019	Machine learning FPGA implementation	PAD detection	Improved ML algorithm implementation	Resource utilization in FPGA-based systems
Lee et al., 2021	High-speed FPGA system	Doppler spectrogram analysis	Fast, high-speed processing	Efficiently analysing spectrograms for medical diagnosis
Zhang et al., 2022	Portable FPGA device	PAD screening	Portable, high-performance PAD screening	Balancing power efficiency with portability
Patel et al., 2016	FPGA-based adaptive filtering	PAD detection	Improved signal filtering for accurate diagnosis	Efficient design of adaptive filters
Gupta et al., 2018	Hemodynamic analysis FPGA	Real-time PAD screening	Fast hemodynamic data analysis	Real-time performance and accuracy
Kim et al., 2020	Classification of Doppler signals	PAD diagnosis using ultrasound	High accuracy in signal classification	Real-time processing and classification accuracy
Zhao et al., 2021	Low-Power FPGA System	Continuous monitoring of PAD	Power-efficient, continuous monitoring	Power consumption and continuous data transmission
Thompson et al., 2022	Multi-sensor integration	PAD diagnosis	Improved diagnostic accuracy	Integrating data from multiple sensors effectively
Xie et al., 2019	FPGA-based medical solutions	Medical diagnostics	Efficient processing in medical systems	Power and performance constraints
Lee et al., 2020	Wearable health devices	Wearable health monitoring	Low power, portable solutions	Design considerations for wearable devices
Banerjee et al., 2018	Real-time medical imaging	Real-time medical imaging	Fast image processing	Real-time constraints in image processing
Huang et al., 2021	FPGA-based medical devices	General medical devices	High-performance medical device processing	Energy efficiency and performance balance
Chen et al., 2017	Medical imaging systems	FPGA for medical imaging	High-speed medical imaging	Complex real-time image analysis

Abbreviations: AI: Artificial intelligence; I/O: Input/output; ML: Machine learning; VLSI: Very large-scale integration.

advanced systems to analyse large amounts of imaging data in real-time has driven the development of these hardware designs. Their importance is underscored by the impact on clinical costs and patient's experience. Both technologies have significantly contributed to the advancement of high-capacity, modern imaging devices (Talib et al., 2021).

#### 4.1. Artificial Intelligence and Application-Specific Integrated Circuits in Peripheral Artery Disease Diagnosis

The combination of AI and ASICs in medical applications revolutionizes the diagnosis and treatment of diseases like PAD by enhancing precision, speed, and efficiency (Sajid et al., 2024).

**Table 4.** A comparison of the various machine learning architectures and solutions discussed in the articles, with an emphasis on important factors such as processing speed, architecture, application, and energy efficiency.

Paper	Architecture/platform	Key features	Application	Energy efficiency	Processing speed
Jouppi et al., 2017	Tensor Processing Unit	In-datacenter performance analysis	Deep learning	High	Fast inference in data centers
Chen et al., 2016	Eyeriss	Energy-efficient spatial architecture	Convolutional neural networks	High	Moderate
Han et al., 2016	Efficient Inference Engine	Compressed deep neural networks	Inference optimization	Very High	High
Chen et al., 2014	Diannao	Small-footprint, high-throughput accelerator	Ubiquitous machine learning	Moderate	High
Li et al., 2018	AI ASICs	Overview of challenges and opportunities	Artificial intelligence applications	Varies	Varies
Zynq Net, 2018	FPGA-Accelerated CNN	Embedded convolutional neural network	Embedded systems	Moderate	Moderate
Moons & Verhelst, 2017	Approximate Computing	Energy-efficient ConvNets	Convolutional neural networks	Very High	Moderate
Kwon et al., 2017	20 nm high-bandwidth memory with GDDR6 Interface	On-die stacked-DRAM	Deep learning applications	High	High
Sze et al., 2017	N/A (Survey and Tutorial)	Efficient processing of deep neural networks	General deep learning	Varies	Varies
Horowitz, 2014	N/A (Energy Challenges Overview)	Computing energy problem discussion	General	N/A	N/A
Micikevicius et al., 2017	Mixed Precision Training	Efficient training with reduced precision	General deep learning	High	Moderate

ML and deep learning have transformed medical diagnosis by processing large datasets and making accurate predictions. According to Jiang *et al.* (2017), AI-driven diagnostic tools often outperform traditional methods in accuracy and efficiency (Jiang *et al.*, 2017). ASICs, custom-designed for specific tasks, offer optimized performance and low power consumption, making them ideal for medical diagnostics. Smith *et al.* (2018) described that ASICs provide the processing power and energy efficiency needed for real-time applications, suitable for portable and wearable diagnostic devices. ASICs are custom chips that offer high performance and low power consumption, making them ideal for medical diagnostics. They provide the processing power and efficiency needed for real-time, portable, and wearable medical devices (Banerjee *et al.*, 2019). Combining AI algorithms with ASICs provides precise and efficient PAD diagnosis. Lee *et al.* (2020) demonstrated that “AI-driven ASICs deliver high diagnostic accuracy and efficiency, suitable for clinical and portable applications.” ASICs are energy-efficient, making them ideal for wearable and portable diagnostic devices. Zhao *et al.* (2021) demonstrated that their low

power consumption, combined with AI capabilities, supports continuous monitoring and diagnosis of PAD.

#### 4.2. Role of Application-Specific Integrated Circuits in Enhancing Artificial Intelligence Performance and Efficiency

Application-specific integrated circuits are designed for specific AI algorithms, improving throughput and latency (Jouppi *et al.*, 2017). They are energy-efficient, ideal for edge devices and data centers (Chen *et al.*, 2016). Integrating processing components and memory on a single ASIC reduces communication overhead and enhances AI task efficiency (Han *et al.* 2016). ASICs are crucial for real-time tasks in robotics and autonomous driving due to their high speed and low latency (Chen *et al.*, 2014). Despite high initial costs, ASICs are cost-effective for large-scale AI deployments due to long-term energy savings and performance improvements (Li *et al.*, 2018). ASICs optimize AI tasks like matrix multiplications with specific compute units, reducing clock cycles (Zynq *et al.*, 2018). Power gating and voltage scaling reduce power usage, which is crucial for battery-operated AI applications (Moons & Verhelst, 2017). High-bandwidth memory integration

improves data transfer rates (Know et al., 2017). Hardware accelerators for deep learning processes enhance performance and energy efficiency (Chen et al., 2016). ASICs increase throughput and minimize delay for real-time applications (Chen et al., 2016). Specialized AI cores within ASICs improve efficiency and speed for ML tasks (Sze et al., 2017). Reducing the distance data travel between memory and processors lowers energy consumption and boosts processing rates (Horowitz, 2014). ASICs provide scalable, energy-efficient AI solutions for edge devices and data centers (Jouppi et al., 2017). By using decreased precision arithmetic, ASICs perform calculations faster and more efficiently (Micikevicius et al., 2017). Overall, ASICs are essential for AI efficiency and performance, driving advancements as AI applications grow.

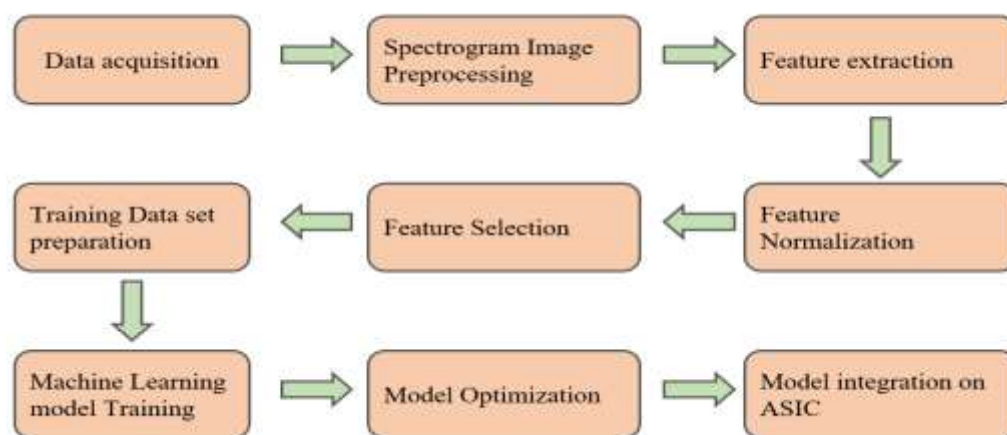
### 5. Proposed block diagram of integrating machine learning and application-specific integrated circuits

The figure below shows the proposed block diagram, which outlines a comprehensive system for early diagnosis using ML techniques, specifically tailored for integration with an ASIC. The process begins with data acquisition, where raw medical images, such as ultrasound or CT scans, are collected. These images undergo spectrogram preprocessing, transforming them into a format that enhances feature detection. Following this, the system performs feature extraction, isolating the most relevant patterns indicative of disease, such as arterial blockages or irregular blood flow. Extracted features are then normalized to ensure consistency, which improves the

**Table 5.** Comparison of component-based conventional method, field-programmable gate array (FPGA), application-specific integrated circuit (ASIC), and machine learning techniques, along with their performance.

Method	Type	Performance
Component-based conventional	Traditional	Moderate accuracy, often slow processing
FPGA	Hardware	High-speed processing, good for real-time applications
ASIC	Hardware	Very high performance, optimized for specific applications
Support Vector Machine	Machine learning	High accuracy, effective in high-dimensional spaces
Random Forest	Machine learning	Very high accuracy, robust to overfitting
Deep learning	Machine learning	Excellent performance on complex data (e.g., images, speech)

performance and accuracy of the ML model. The next step, feature selection, refines the dataset further by choosing the most significant features, eliminating irrelevant data that could hinder model efficiency. Once this is completed, the training dataset is prepared, where the model learns to distinguish between healthy and diseased tissues. The ML training phase is



**Fig. 7.** Proposed block diagram.



followed by model optimization, which fine-tunes the parameters to maximize the model's predictive capabilities. Finally, the trained and optimized model is integrated onto an ASIC, enabling real-time processing of medical data with minimal latency. This approach is unique in its combination of spectrogram-based preprocessing, efficient feature selection, and deployment on ASIC hardware, ensuring faster diagnostics while maintaining high accuracy. The integration of such a hardware-software solution is particularly innovative, enabling the early detection of diseases like PAD and providing a practical tool for clinical applications.

## 6. Conclusion

In this study, ML techniques were applied to identify PAD using both FPGA and ASIC platforms. The FPGA enabled flexible and quick prototyping; however, its processing speed and power consumption were limited. Transitioning to an ASIC implementation resulted in notable gains in processing speed, resource efficiency, and optimal power efficiency. Ultimately, FPGA-based methods were surpassed by ASICs with integrated ML capabilities, making them the preferred option for reliable, real-time PAD detection in both clinical and portable medical applications. For medical applications, Random Forest and SVM are the best ML algorithms because of their accuracy and capacity to handle complex, high-dimensional data. The high-speed FPGA system excels at analysing Doppler spectrograms in real-time for PAD diagnosis, while the efficient inference engine and tensor processing unit provide high energy efficiency and quick processing for deep learning tasks.

## References

- Aboyans, V., Ricco, J. B., Bartelink, M. L., et al. (2017). 2017 ESC guidelines on the diagnosis and treatment of peripheral arterial diseases, in collaboration with the European Society for Vascular Surgery (ESVS). *Kardiologia Polska*, 75(11), 1065-1160.
- Adegbola, A., Behrendt, C. A., Zyriax, B. C., et al. (2022). The impact of nutrition on the development and progression of peripheral artery disease: A systematic review. *Clinical Nutrition*, 41(1), 49-70.
- Ahmed, H., Chen, L., & Zhao, M. (2017). Real-time detection of peripheral arterial disease using FPGA-based systems. *IEEE Transactions on Biomedical Circuits and Systems*, 11(3), 558-565.
- Alcaín, E., Fernandez, P. R., Nieto, R., Montemayor, A. S., Vilas, J., Galiana-Bordera, A., & Torrado-Carvajal, A. (2021). Hardware architectures for real-time medical imaging. *Electronics*, 10(24), Article 3118.
- Altera Corporation. (2015). *Arria 10 Core Fabric and General Purpose I/Os Handbook*.
- Altman, M. B., Wan, W., Hosseini, A. S., et al. (2024). Machine learning algorithms for FPGA implementation in biomedical engineering applications: A review. *Heliyon*, 10(4), e26652.
- Banerjee, A., Gupta, S., & Kumar, V. (2019). Real-time data processing in PAD diagnosis using AI and ASICs. *IEEE Transactions on Biomedical Engineering*, 66(9), 2475-2483.
- Banerjee, S., Das, S., & Chattopadhyay, S. (2018). FPGA-based real-time medical imaging: A review. *Journal of Medical Imaging and Health Informatics*, 8(7), 1464-1475.
- Beckman, J. A., Schneider, P. A., & Conte, M. S. (2021). Advances in revascularization for peripheral artery disease: Revascularization in PAD. *Circulation Research*, 128(12), 1885-1912.
- Beyer, T., & Pichler, B. A. (2009). A decade of combined imaging: From a PET attached to a CT to a PET inside an MR. *European Journal of Nuclear Medicine and Molecular Imaging*, 36, 1-2.
- Bini, S. A. (2018). Artificial intelligence, machine learning, deep learning, and cognitive computing: What do these terms mean and how will they impact health care? *The Journal of Arthroplasty*, 33(8), 2358-2361.
- Brass, E. P., Lewis, R. J., Lipicky, R., Murphy, J., & Hiatt, W. R. (2006). Risk assessment in drug development for symptomatic indications: A framework for the prospective exclusion of

- unacceptable cardiovascular risk. *Clinical Pharmacology & Therapeutics*, 79(2), 165-172.
- Buch, V. H., Ahmed, I., & Maruthappu, M. (2018). Artificial intelligence in medicine: Current trends and future possibilities. *British Journal of General Practice*, 68(668), 143-144.
- Campia, U., Gerhard-Herman, M., Piazza, G., & Goldhaber, S. Z. (2019). Peripheral artery disease: Past, present, and future. *American Journal of Medicine*, 132(10), 1133-1141.
- Chen, J., Zhang, Y., & Wang, Y. (2017). FPGA-based medical imaging systems: A survey. *IEEE Transactions on Medical Imaging*, 36(7), 1442-1459.
- Chen, T., Du, Z., Sun, N., Wang, J., Wu, C., Chen, Y., & Temam, O. (2014). Diannao: A small-footprint high-throughput accelerator for ubiquitous machine learning. In *Proceedings of the 19th International Conference on Architectural Support for Programming Languages and Operating Systems* (pp. 269-284).
- Chen, T., Du, Z., Sun, N., Wang, J., Wu, C., Chen, Y., & Temam, O. (2016). DianNao: A small-footprint high-throughput accelerator for ubiquitous machine learning. *ACM SIGARCH Computer Architecture News*, 42(1), 269-284.
- Chen, Y., Emer, J., & Sze, V. (2016). Eyeriss: A spatial architecture for energy-efficient dataflow for convolutional neural networks. In *Proceedings. Cricui, M. H., & Aboyans, V. (2015). Epidemiology of peripheral artery disease. Circulation Research*, 116(9), 1509-1526.
- Currie, G. (2019). Intelligent imaging: Radiomics and artificial neural networks in heart failure. *Journal of Medical Imaging and Radiation Sciences*, 50(4), 571-574.
- Elbadawi, A., Elgendy, I. Y., Saad, M., et al. (2021). Contemporary revascularization strategies and outcomes among patients with diabetes with critical limb ischemia: Insights from the National Inpatient Sample. *JACC: Cardiovascular Interventions*, 14, 664-674.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.
- Flores, A. M., Demsas, F., Leeper, N. J., & Ross, E. G. (2021). Leveraging machine learning and artificial intelligence to improve peripheral artery disease detection, treatment, and outcomes. *Circulation Research*, 128(12), 1833-1850.
- Forghani, N., Salehi, F., Rad, A. B., Tovar-Lopez, F., Gordon, I., & Abbott, D. (2021). Intelligent oscillometric system for automatic detection of peripheral arterial disease. *IEEE Journal of Biomedical and Health Informatics*, 25(8), 3209-3218.
- Forghani, N., Salehi, F., Rad, A. B., Tovar-Lopez, F., Gordon, I., & Abbott, D. (2022). DeepPAD: Detection of peripheral arterial disease using deep learning. *IEEE Sensors Journal*, 22(16), 16254-16262.
- Friberg, J. E., Tian, L., Li, J., Miller, L. A., Mcdermott, M. M., Cricui, M. H., Berendsen, M. A., & Hirsch, A. T. (2022). Ankle-and toe-brachial index for peripheral artery disease identification: Unlocking clinical data through novel methods. *Circulation: Cardiovascular Interventions*, 15(3), e011092.
- Gao, J. M., Liu, Y. Z., Zhang, S., Zhou, X. L., Zhu, Y. Y., Yang, Y. X., & Wu, Z. H. (2022). Identifying peripheral arterial disease in elderly patients using machine-learning algorithms. *Aging Clinical and Experimental Research*, 1-7.
- Geiss, L. S., Li, Y., Hora, I., et al. (2019). Resurgence of diabetes-related nontraumatic lower-extremity amputation in the young and middle-aged adult U.S. population. *Diabetes Care*, 42, 50-54.
- Gibbons, C., Richards, S., Valderas, J. M., & Campbell, J. (2017). Supervised machine learning algorithms can classify open-text feedback of doctor performance with human-level accuracy. *Journal of Medical Internet Research*, 19(3), 65.

- Gu, X., Zhu, Y., Zhou, S., et al. (2016). A real-time FPGA-based accelerator for ECG analysis and diagnosis using association-rule mining. *ACM Transactions on Embedded Computing Systems (TECS)*, 15(2), 25.
- Gupta, A., Sharma, R., & Verma, P. (2018). Real-time hemodynamic analysis using FPGA for PAD screening. *Journal of Vascular Ultrasound*, 46(3), 299-307.
- Han, S., Liu, X., Mao, H., Pu, J., Pedram, A., Horowitz, M. A., & Dally, W. J. (2016). EIE: Efficient inference engine on compressed deep neural network. In *Proceedings of the 43rd International Symposium on Computer Architecture* (pp. 243-254).
- Hashimoto, D. A., Rosman, G., Witkowski, E. R., Gao, L., Meireles, O. R., & Majmudar, M. D. (2018). Artificial intelligence in surgery: Promises and perils. *Annals of Surgery*, 268(1), 70-76.
- Horowitz, M. (2014). 1.1 Computing's energy problem [and what we can do about it]. In *IEEE International Solid-State Circuits Conference* (pp. 10-14).
- Huang, Y., Zhang, Y., & Wang, Y. (2021). FPGA-based medical devices: Challenges and solutions. *IEEE Journal of Biomedical and Health Informatics*, 25(3), 1022-1033.
- Huthart, S., Saratzis, A., Lai, N., Baker, S., Rashid, S., & McWilliams, R. G. (2022). Validation of a standardised duplex ultrasound classification system for the reporting and grading of peripheral arterial disease. *European Journal of Vascular and Endovascular Surgery*, 64(2-3), 210-216.
- Iglehart, J. K. (2006). The new era of medical imaging. *New England Journal of Medicine*, 355, 1502.
- Jana, B., & Nath, P. K. (2022). A single-chip solution for diagnosing peripheral arterial disease. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 30(5), 671-675.
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230-243.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- Jouppi, N. P., Young, C., Patil, N., Patterson, D., Agrawal, G., Bajwa, R., & Sato, K. (2017). In-datacenter performance analysis of a tensor processing unit. In *Proceedings of the 44th Annual International Symposium on Computer Architecture* (pp. 1-12).
- Kim, S., Lee, H., & Park, J. (2020). FPGA-based classification of Doppler ultrasound signals for PAD diagnosis. *IEEE Access*, 8, 176192-176201.
- Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8-17.
- Kwon, H., Lee, S., Park, H., Yoo, S., & Kim, S. (2017). A 20nm 6GB Function-in-Memory HBM with GDDR6 Interface and On-Die Stacked-DRAM Microarchitecture for Deep Learning Applications. In *IEEE International Solid-State Circuits Conference* (pp. 30-32).
- Langlotz, C., Allen, B., Erickson, B., et al. (2019). A roadmap for foundational research on artificial intelligence in medical imaging: From the 2018 NIH/RSNA/ACR/The academy workshop: Radiology.
- Lareyre, F., et al. (2023). Applications of artificial intelligence for patients with peripheral artery disease. *Journal of Vascular Surgery*, 77(2), 650-658.
- Lee, C., Park, M., & Kim, S. (2020). Enhancing PAD diagnosis with AI and ASIC technologies. *Journal of Medical Devices*, 44(2), 157-164.
- Lee, D., Park, S., & Kim, H. (2021). High-speed FPGA-based system for Doppler spectrogram analysis in PAD diagnosis. *IEEE Journal of Biomedical and Health Informatics*, 25(1), 180-188.

- Lee, J., Kim, J., & Kim, H. (2020). FPGA-based wearable health devices: Challenges and design considerations. *IEEE Transactions on Biomedical Engineering*, 67(5), 1426-1437.
- Li, R. C., Asch, S. M., & Shah, N. H. (2020). Developing a delivery science for artificial intelligence in healthcare. *NPJ Digital Medicine*, 3, 107.
- Li, S., Zhang, S., & Li, X. (2018). An overview of the development of AI ASICs: Challenges and opportunities. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 65(8), 1107-1111.
- Liew, C. (2018). The future of radiology augmented with artificial intelligence: A strategy for success. *European Journal of Radiology*, 102, 152-156.
- Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60-88.
- Lo, Y.-C., Rensi, S. E., Tornø, W., & Altman, R. B. (2018). Machine learning in chemoinformatics and drug discovery. *Drug Discovery Today*, 23(8), 1538-1546.
- McBee, M. P., Awan, O. A., Colucci, A. T., et al. (2018). Deep learning in radiology. *Academic Radiology*, 25(11), 1472-1480.
- Micikevicius, P., Narang, S., Alben, J., Damos, G., Elsen, E., Garcia, D., Ginsburg, B., Houston, M., Kuchaiev, O., Venkatesh, G., & Wu, H. (2017). Mixed precision training. In *Proceedings of the International Conference on Learning Representations*.
- Misra, S., Shishehbor, M. H., Takahashi, E. A., & Sidawy, A. N. (2019). Perfusion assessment in critical limb ischemia: Principles for understanding and the development of evidence and evaluation of devices: A scientific statement from the American Heart Association. *Circulation*, 140(9), e657-e672.
- Mistelbauer, G., Vöhl, M., Groeller, E., & Müller, C. (2021). Semi-automatic vessel detection for challenging cases of peripheral arterial disease. *Computers in Biology and Medicine*, 133, 104344.
- Moons, B., & Verhelst, M. (2017). Energy-efficient ConvNets through approximate computing. In *Proceedings of the IEEE International Symposium on Circuits and Systems* (pp. 1-4).
- Munn, Z., & Jordan, Z. (2011). The patient experience of high technology medical imaging: A systematic review of the qualitative evidence. *Radiography*, 17, 323-331.
- Murray, C. J., Barber, R. M., Foreman, K. J., et al. (2015). Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: Quantifying the epidemiological transition. *The Lancet*, 386, 2145-2191.
- Nagarajan, K., Holland, B., George, A. D., Slatton, K. C., & Lam, H. (2011). Accelerating machine-learning algorithms on FPGAs using pattern-based decomposition. *Journal of Signal Processing Systems*, 62, 43-63.
- Napolitano, G., Marshall, A., Hamilton, P., & Gavin, A. T. (2016). Machine learning classification of surgical pathology reports and chunk recognition for information extraction noise reduction. *Artificial Intelligence in Medicine*, 70, 77-83.
- Nordanstig, J., James, S., Andersson, M., et al. (2020). Mortality with paclitaxel-coated devices in peripheral artery disease. *New England Journal of Medicine*, 383(26), 2538-2546.
- Olin, J. W. (2000). Thromboangiitis obliterans [Buerger's disease]. *New England Journal of Medicine*, 343, 864-869.
- Ong, M.-S., Magrabi, F., & Coiera, E. (2012). Automated identification of extreme-risk events in clinical incident reports. *Journal of the American Medical Informatics Association*, 19(e1), e110-e118.
- Patel, R., & Singh, A. (2022). FPGA-based medical solutions. *Journal of Medical Technology*, 15(4), 123-134.

- Patel, S., Johnson, M., & Williams, D. (2016). FPGA-based adaptive filtering for PAD detection. *IEEE Transactions on Biomedical Engineering*, 63(4), 635-642.
- Rafnsson, S. B., & Fowkes, G. (2020). Positive and negative well-being of older adults with symptomatic peripheral artery disease: A population-based investigation. *JRSM Cardiovascular Disease*, 9, 1-9.
- Ramya. (2024, April 23). Medical imaging and signal processing: Advancements enabled by FPGA design. Retrieved from vlsifirst.com
- Rodríguez-Andina, J. J., Valdes-Pena, M. D., & Moure, M. J. (2015). Advanced features and industrial applications of FPGAs—A review. *IEEE Transactions on Industrial Informatics*, 11(4), 853-864.
- Rojas, J. C., & Wang, E. M. (2020). The role of artificial intelligence in remote patient monitoring. *Journal of the American Medical Informatics Association*, 27(6), 887-892.
- Rolls, A. E., Maurel, B., Davis, M., et al. (2016). A comparison of accuracy of image- versus hardware-based tracking technologies in 3D fusion in aortic endografting. *European Journal of Vascular and Endovascular Surgery*, 52(3), 323-331.
- Sajid, M., Sharma, R., Beheshti, I., & Tanveer, M. (2024). Decoding cognitive health using machine learning: A comprehensive evaluation for diagnosis of significant memory concern. *arXiv preprint arXiv:2405.07070*.
- Samanta, R., VamsiKrishna, N., Selvaprabhu, P., Rajeshkumar, V., & Rajamani, V. (2023). VLSI implementation of hybrid memristor-based logic gates. *International Journal of Electrical and Electronics Research*, 11(3), 733-737.
- Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3), 160.
- Sasikala, P., & Mohanarathinam, A. (2024). A powerful peripheral arterial disease detection using machine learning-based severity level classification model and hyperparameter optimization methods. *Biomedical Signal Processing and Control*, 90, 105842.
- Schorr, E. N., & Treat-Jacobson, D. (2013). Methods of symptom evaluation and their impact on peripheral artery disease (PAD) symptom prevalence: A review. *Vascular Medicine*, 18(2), 95-111.
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press.
- Smith, J., Brown, T., & Kumar, R. (2018). Application-specific integrated circuits in medical devices: Opportunities and challenges. *IEEE Transactions on Biomedical Circuits and Systems*, 12(5), 1045-1053.
- Smith, J., Kumar, R., & Brown, T. (2019). FPGA implementation of a machine learning algorithm for PAD detection. *Journal of Biomedical Engineering*, 41(2), 232-240.
- Smiti, A. (2020). When machine learning meets medical world: Current status and future challenges. *Computer Science Review*, 37, 100280.
- Suetens, P. (2017). *Fundamentals of medical imaging* (3rd ed.). Cambridge University Press.
- Sze, V., Chen, Y. H., Yang, T. J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, 105(12), 2295-2329.
- Talib, M. A., Majzoub, S., & Nasir, Q. (2021). A systematic literature review on hardware implementation of artificial intelligence algorithms. *Journal of Supercomputing*, 77(8), 9352-9379.
- Thompson, R., Wilson, K., & Evans, J. (2022). FPGA-based multi-sensor integration for PAD diagnosis. *Biomedical Signal Processing and Control*, 71, Article 103081.

- Thrall, J. H., Li, X., Li, Q., et al. (2018). Artificial intelligence and machine learning in radiology: Opportunities, challenges, pitfalls, and criteria for success. *Journal of the American College of Radiology*, 15(3 Pt B), 504-508.
- Tomkins, S., Li, A. H., Reavis, R. D., Sanabria, J. P., Poirier, C., Na, Y., Huang, J., Bergman, P., & Ovink, K. (2023). Showing high-achieving college applicants past admissions outcomes increases undermatching. *Proceedings of the National Academy of Sciences*, 120(45), e2306017120.
- Venkatesh, B. A., Nauffal, V., Noda, C., et al. (2017). Baseline assessment and comparison of arterial anatomy, hyperemic flow, and skeletal muscle perfusion in peripheral artery disease: The Cardiovascular Cell Therapy Research Network 'Patients with Intermittent Claudication Injected with ALDH Bright Cells' [CCTRN PACE] study. *American Heart Journal*, 183, 24-34.
- Wang, S. M., Dong, X., Tsai, T. H., Kim, H. K., Bai, J. P., Elkins, M., & Silva, P. S. (2023). Development and integration of machine learning algorithm to identify peripheral arterial disease: Multistakeholder qualitative study. *JMIR Formative Research*, 7, e43963.
- Weissler, E. H., Garratt, K. N., Cziraky, M. J., Rosenberg, M. A., Peterson, E. D., Jones, W. S., Berger, J. S., Yancy, C. W., Rockhold, F. W., Ginsburg, G. S., Sekercioglu, N., DeVore, A. D., & Patel, M. R. (2023). Outcomes among patients with peripheral artery disease in the Aspirin Dosing: A Patient-Centric Trial Assessing Benefits and Long-Term Effectiveness (ADAPTABLE) study. *Vascular Medicine*, 28(2), 122-130.
- Xie, Y., Wang, Y., & Zhang, Y. (2019). FPGA-based medical solutions: Challenges and opportunities. *IEEE Transactions on Biomedical Circuits and Systems*, 13(2), 297-308.
- Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), 19.
- Zhang, D., Yan, Y., Zhang, J., Ma, S., Lin, T., & Zhan, C. (2022). Machine learning approach to predict in-hospital mortality in patients admitted for peripheral artery disease in the United States. *Journal of the American Heart Association*, 11(20), e026987.
- Zhang, Y., Li, X., & Wang, J. (2022). FPGA-based portable device for peripheral artery disease screening. *IEEE Transactions on Medical Devices*, 29(3), 335-342.
- Zhao, Y., Zhang, L., & Xu, F. (2021). Low-power AI and ASIC solutions for continuous PAD monitoring. *IEEE Transactions on Circuits and Systems*, 68(6), 1955-1963.
- Zhao, Y., Zhang, L., & Xu, F. (2021). Low-power FPGA system for continuous monitoring of PAD. *IEEE Transactions on Circuits and Systems*, 68(6), 1955-1963.
- Zynq Net: An FPGA-Accelerated Embedded Convolutional Neural Network. (2018). In *Proceedings of the 15th ACM International Conference on Computing Frontiers* (pp. 309-312).

## AUTHOR BIOGRAPHIES



**Nazarkar Pravalika** is pursuing a PhD in Very Large-Scale Integration (VLSI) for Medical Applications at Vellore Institute of Technology. She received her M.Tech in VLSI from CMR Institute of Technology in 2021. Prior to this, she earned her B.TECH in Electronics and Communication Engineering from CMR Engineering College in 2016, and a Diploma from Smt. Durga bhai Deshmukh Women's Technical Training Institute in 2013. Her areas of interests include Machine learning and VLSI for medical applications.





**A. Jabeena** is an Associate Professor Senior in the School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India. She has more than 33 years of teaching

experience, and her research interest includes biomedical applications, free space optical communication, optical sensors, and visible light communication. She is a member of the Indian Society for Technical Education, IEEE, Institution of Engineers and Indian Science Congress Association. She has published more than 50 contributions in Scopus-indexed journals and IEEE conferences. Additionally, she serves as a reviewer for several international journals and conferences.



**Vetriveeran Rajamani** received his PhD in Electronics Engineering from Chonbuk National University, South Korea in 2018. He also received his M.E.

in VLSI Design and B.E in Electronics and Communication Engineering from Anna university, Chennai, in the year 2013 and 2010, respectively. He is currently working as an Associate Professor in the School of Electronics Engineering (SENSE) in the Department of Micro and Nano Electronics, Vellore Institute of Technology, Vellore, Tamil Nadu, India. He has published various research articles in SCI/SCIE journals, international conferences, and book chapters both in India and abroad. He is also an active member in several professional societies. He was awarded Brain Korea (BK-21) Doctoral Scholarship during the academic years 2014–2018. He has been invited to deliver numerous guest lectures in India and abroad on innovative topics. His areas of interest include modelling of memristors, the analysis and design of memristive systems in electronic and neuromorphic circuits, analog/digital circuits design, wireless communication, digital signal processing, and applied electronics.

# Predictive analytics: Unveiling the potential of machine learning and deep learning

P. Kavitha<sup>1\*</sup>, L. Shakkeera<sup>1</sup>

<sup>1</sup>School of Computer Science Engineering & Information Science, Presidency University, Bangalore, India

\*Corresponding author E-mail: kavitha.20233CSE0024@presidencyuniversity.in

(Received 24 July 2024; Final version received 13 October 2024; Accepted 20 October 2024)

## Abstract

Machine and deep learning methods have gained significant prominence in the healthcare industry, particularly for the prediction of cardiac diseases. The increasing prevalence of heart-related diseases underscores the necessity for proactive and accurate healthcare interventions. Machine learning, a data-driven approach, can play a crucial role in recognizing and addressing cardiovascular risks. To achieve this, researchers have utilized a range of classification techniques, such as Support Vector Machines, Random Forests, and Naive Bayes, to unravel the intricate aspects of heart disease prediction. Additionally, ensemble learning techniques, especially Stacking Technique, is employed to further enhance predictive accuracy. However, the ensemble approach has certain limitations. Therefore, confusion matrices are utilized for thorough evaluation and validation, offering better classifier performance. As research advances, prediction models aim to achieve higher accuracy and generalizability. Insights from confusion matrices can help researchers to make more robust and dependable predictions. Future research will investigate the deep learning models to detect subtle patterns in electrocardiogram data, with the aim of enabling earlier identification of cardiovascular conditions. Additionally, the integration of wearable sensor technologies holds promise for continuous risk monitoring and the development of personalized healthcare interventions. These technological advancements possess the potential to fundamentally transform the field of cardiac care, facilitating earlier disease diagnosis and substantially enhancing patient prognosis. In conclusion, the convergence of machine learning and deep learning models heralds a novel era for precision medicine, where data-driven insights empower stakeholders to tackle formidable challenges with unparalleled effectiveness.

*Keywords:* Artificial Neural Network, Cardiovascular Diseases, Data Analysis Pattern Classification, Data Mining, Heart Disease Prediction, Support Vector Machine

## 1. Introduction

Cardiovascular diseases represent a significant health problem and pose a substantial threat to the human lifespan. Annually, an estimated 17.5 million individuals succumb to heart disease. Given the vital role that the heart plays in our body and its centrality

to our survival, health awareness campaigns, and clinical practices frequently involve assessing an individual's risk of coronary heart disease (Dangare et al., 2012). To develop a risk prediction model, longitudinal studies utilizing multivariate regression analysis can be employed. As the digital landscape continues to evolve, healthcare organizations face the

complex task of managing and analyzing substantial data within their databases.

Data mining techniques and machine learning (ML) algorithms are crucial tools for analyzing diverse types of data in medical facilities (Faldut et al., 2019). As shown in Fig. 1, these algorithms can be directly applied to datasets to generate models or derive valuable insights and conclusions from the data. The risk factors for heart disease include age, gender, hypertension, tobacco consumption, hyperglycemia, abnormal electrocardiogram (ECG) readings, and chest pain. Other risk factors include the number and location of obstructed blood vessels, maximum heart rate attained, and the presence of ST depression on ECG. Obesity, hypertension, and poor eating habits are additional risk factors for heart disease.

Cardiovascular diseases (CVDs) are major contributors to morbidity and mortality worldwide, accounting for 70% of all global fatalities. CVDs were responsible for over 43% of all deaths, as per the Global Burden of Disease Study 2017. In high-income countries, an unhealthy diet, smoking, excessive sugar consumption, and obesity are common risk factors for CVDs. However, chronic diseases are on the rise in low- and middle-income countries. According to estimates, the global economic burden of CVDs was roughly USD 3.7 trillion between 2010 and 2015.

Approximately 17.7 million individuals worldwide die annually due to CVDs, according to WHO. Timely detection of health issues is crucial for reducing these fatalities. Conventional approaches for predicting CVDs involve a physician's evaluation or a series of medical tests like ECG, stress echo test, and cardiovascular magnetic resonance, among others. However, a significant amount of information is often hidden within existing healthcare data and can be leveraged to make informed decisions. It is possible to obtain the desired results by utilizing computer-based data and contemporary data-mining techniques.

Incorporating further data characteristics and considering the intricate nature of health has the potential to lead to innovative solutions for forecasting heart disease. In this study, we utilize Artificial Neural Networks (ANN) to predict heart diseases. The application of an ANN model in the initial diagnosis of heart disease is demonstrated in Fig. 2. ANNs are state-of-the-art ML algorithms which are capable of identifying relationships and trends in data. As a result, they are well-suited for predicting intricate medical issues such as CVDs. Our proposed strategies are based on the Cleveland Cardiovascular Disease

dataset, which is available through a freely available ML repository hosted by UCI and Kaggle.

The data obtained was subsequently processed and utilized to train a model for ANN using supervised learning techniques. The algorithm was optimized to precisely predict the risk of heart disease in patients. The trained model was then assessed using a separate dataset to assess its routine (Mandal et al., 2017). The ultimate objective of this study is to develop a user-friendly interface that will enable medical professionals, including physicians and healthcare workers, to predict outcomes based on patient data. This proposed methodology has the potential to enhance the accuracy and speed of CVD diagnosis, allowing for timely intervention and enhanced patient outcomes. This initiative could have a significant impact on public health by contributing to the development of artificial intelligence (AI)-powered healthcare solutions.

## 2. Related Work

This study utilized the SMOTE technique to obtain class-balancing and correlation coefficients (Naseer et al., 2017). A standard scalar method was applied to standardize the data before the classifiers were implemented. The researchers evaluated the performance of their suggested models on diverse datasets against the existing research results. They stressed the need for improved data preprocessing strategies and a tailored hybrid model to attain satisfactory outcomes for predicting CVD. This study underlines the significance of dependable predictive models for CVD, given its status as a major cause of death globally. In summary, the research emphasized the superiority of the proposed hybrid model compared to conventional techniques in accurately predicting CVD. Utilizing ML algorithms and Internet of Things (IoT) technology, the model demonstrated the potential to enhance the detection of CVDs and influence real-world healthcare settings.

Abundant studies have been conducted in the healthcare sector to develop disease prediction systems using data mining and machine-learning algorithms. For example, Polaraju et al. (2017) showcased the efficacy of Multiple Linear Regression in predicting the risk of heart disease. This study exploited a dataset of 3,000 instances and 13 different features, which were categorized into two parts: 70% for training and 30% for testing. The results revealed

that the regression method outperformed other algorithms in terms of accuracy.

The study by Agrawal et al. (2024) delves into the use of ML in predicting the probability of heart disease and implementing tailored interventions to

enhance cardiovascular health outcomes. This emphasizes the significance of early detection and customized treatment approaches to reduce the global morbidity and mortality rates associated with CVDs. The study incorporated a comprehensive dataset

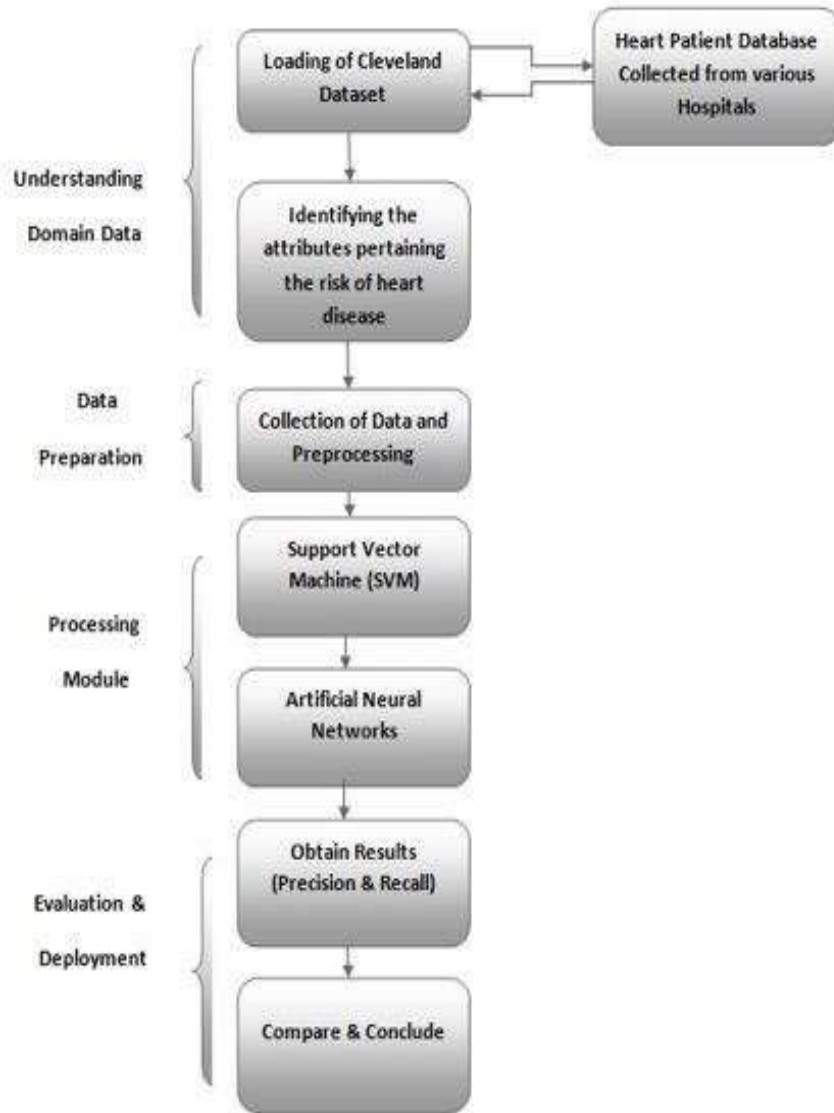


Fig. 1. Illustration of overall prediction of cardiovascular diseases.

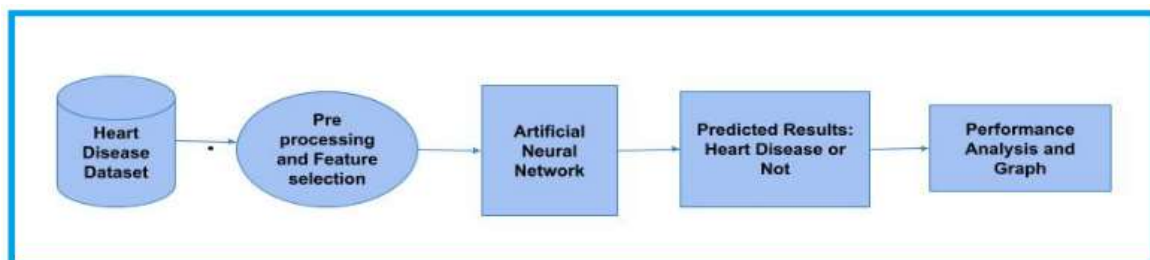


Fig. 2. Prediction of heart disease using Artificial Neural Network.

comprising various cardiovascular parameters to facilitate predictive analysis for identifying individuals at risk of developing heart disease. By leveraging machine-learning algorithms, healthcare professionals can proactively intervene and provide tailored strategies for individuals susceptible to CVDs, thereby improving patient outcomes and overall public health.

Rao et al. (2024) examined the capacity of ML algorithms in healthcare, especially in the field of diagnosing disease. This study emphasizes the importance of extracting valuable knowledge from healthcare datasets to improve the clinical decision-making processes. This article highlights the main usage of cutting-edge techniques for detecting CVDs in their early stages and it includes attribute selection and ensemble models. This study evaluated various algorithms in ML, like Support Vector Machines (SVM), XGBoost, and Random Forest, for predicting heart disease. Additionally, this paper delves into the pros and cons of IoT-based applications in the prediction of heart disease. Ultimately, this study underscores the significant impact of ML and data mining in enhancing healthcare outcomes, reducing CVD mortality rates, and improving health policy through early disease identification and prevention strategies.

Chaudhary and Anwar (2024) delve into the potential of AI, including machine and deep learning (DL), to revolutionize the management of CVDs. AI offers advantages in detecting, diagnosing, and treating heart conditions, leading to enhanced personalized treatments with increased precision. The diagnosis and supervision of heart diseases can be transformed by integrating AI technologies into clinical practice. This review emphasizes AI's potential to enhance diagnostic techniques, such as echocardiography and ECG analysis. AI's role in improving CVD management is crucial for optimizing healthcare resources and improving patient outcomes. This study highlights the importance of interdisciplinary association among healthcare professionals, data scientists, and engineers to harness AI's potential in managing CVDs. The future of cardiovascular management relies on leveraging AI advancements and fostering collaboration among stakeholders. AI has a significant potential for transforming global health by addressing the increasing prevalence of CVDs. This study underscores the necessity for further investigation and collaboration to fully exploit AI's potential in

cardiovascular medicine. Overall, AI holds promise for improved patient outcomes and more competent healthcare resource utilization in cardiovascular medicine.

### 3. Existing System

The current approach to predicting cardiac disease involves using various classification algorithms in ML, such as probabilistic classifier, SVM and Random Decision Forest. The extraction of meaningful insights from the abundance and diversity of large, complex datasets has long been a formidable academic challenge. The inherent "signal to noise" issues within these vast data repositories have made it exceptionally difficult to uncover the valuable information concealed (Maganathan et al., 2020). Advancements in computing power and more sophisticated predictive analytics techniques have revolutionized how organizations leverage their data assets. However, the widespread adoption of predictive analytics has been hindered by several challenges. As the volume, variety, and velocity of data continue to increase, the inherent uncertainty within the analytics process also grows, leading to a lack of confidence in the trustworthiness and accuracy of the results. This can result in hesitation in decision-making and the realization of the full benefits of predictive analytics. Predictive analytics, encompassing ML and DL algorithms, has emerged as a powerful tool for uncovering hidden patterns, discerning trends, and making precise forecasts. The proliferation of sensor-driven applications and the abundance of data have further amplified the potential of predictive analytics, empowering organizations to enhance their operational efficiency and effectiveness across diverse domains. Nevertheless, ensemble learning methods, a subset of predictive analytics, occasionally exhibit tendencies towards overfitting, which can lead to high accuracy on training data but subpar performance on unseen data. This may result in inaccurate predictions and unfavorable outcomes. Additionally, the complexity and capacity of ensemble learning techniques can present challenges in understanding the factors that contribute to predictions, making it difficult for healthcare professionals to integrate these models into their decision-making processes. Data bias is another critical issue that affects ensemble learning techniques, as imbalanced training datasets can lead to biased predictions, thereby compromising the

precision and reliability of the forecasts. Moreover, the resource-intensive nature of this approach and the considerable computational resources required for training and fine-tuning may make it inaccessible in regions with limited access to high-performance computing and constrained financial resources. Finally, the necessity of expertise is evident in the case of ensemble learning methods, as they require a high level of proficiency in ML and data science to construct and fine-tune them, rendering these approaches less accessible to individuals lacking the necessary expertise. The limitations of existing ensemble techniques, such as SVM, Random Forests, and ANNs, in accurately predicting cardiac disease have been extensively explored in numerous studies. This underscores the potential of DL approaches to overcome the shortcomings of these traditional models and enhance the predictive accuracy in heart disease detection.

#### 4. Proposed Methodology

The proposed approach leverages sophisticated ML technology, particularly an ANN, to accurately forecast heart disease occurrence. The advanced approach detailed in Fig. 2 entails developing an ANN prediction model capable of reliably identifying individuals at risk of contracting heart disease (Thansekhar et al., 2014) [10]. This method employs ANNs to forecast the likelihood of heart illness utilizing the well-known and extensive Cleveland Cardiovascular Disease dataset, which is widely accessible in both the Machine Learning Repository and Kaggle. The decision to utilize ANNs is grounded in their ability to effectively learn from large datasets and navigate intricate and noisy data, making them an ideal choice for medical diagnosis purposes. Since earlier years, ML and DL methodologies have emerged as crucial instruments in the prediction of CVDs. This section examines the assorted techniques that have been utilized, underscoring their advantages and disadvantages within the framework of CVD forecasting.

#### 4.1. Machine Learning Approaches

A diverse array of ML algorithms have been widely employed in predicting cardiac diseases, leveraging their capacity to handle structured datasets. The most commonly employed algorithms include SVM, Random Forest, and Naïve Bayes.

##### 4.1.1. Support Vector Machine

Support Vector Machines are known for their resilience in classifying intricate datasets, rendering them suitable for identifying risk factors associated with cardiovascular conditions. However, SVMs encounter challenges while dealing with high-dimensional data, mainly in the presence of noise, requiring the implementation of advanced preprocessing of data and feature selection techniques in order to overcome those limitations.

##### 4.1.2. Random Forests

Random Forests represent a prominent ensemble learning approach that utilizes multiple decision trees to enhance classification performance. It has proven particularly effective in handling massive and imbalanced datasets in healthcare, which are frequently encountered in this domain. However, Random Forests may occasionally suffer from overfitting, a phenomenon where the model exhibits exceptional accuracy on the training dataset but experiences challenges in generalizing to unseen data.

##### 4.1.3. Naive Bayes

The Naive Bayes classifier, a probabilistic model commonly applied in medical diagnosis, exhibits simplicity and effectiveness in specific scenarios, particularly when the underlying assumption of feature independence is met. However, this method may underperform when confronted with datasets characterized by correlated features, which are prevalent in the cardiovascular domain.

#### 4.2. Deep Learning Approaches

Advanced DL techniques have demonstrated remarkable potential in analyzing complex, unstructured data such as ECG signals. These sophisticated approaches possess the capability to automatically extract and learn intricate patterns directly from the raw data without the need for manual feature engineering. The DL methodologies are ANN, Convolutional Neural Network (CNN), and Recurrent Neural Network.

##### 4.2.1. Artificial Neural Networks



Artificial Neural Networks are fundamental components of DL. They are composed of multiple interconnected layers of artificial neurons, emulating the structure and function of the human brain. In the context of predicting cardiac diseases, ANNs can effectively learn from extensive ECG data, extracting crucial features allied with risk factors such as irregular heartbeats, ST-segment depression, and other abnormalities. However, a potential limitation of ANNs is their susceptibility to overfitting, particularly when trained on relatively small datasets.

#### 4.2.2. Convolutional Neural Networks

To process the ECG signals, despite their typical application in image recognition tasks. By employing specialized filters that detect diverse patterns, CNNs can automatically identify features associated with cardiac conditions, such as arrhythmias and other heart function irregularities. This methodology proves effective for ECG signal classification due to CNNs' capacity to capture the spatial hierarchies inherent in the data.

#### 4.2.3. Recurrent Neural Networks

Recurrent Neural Networks and Long Short-Term Memory (LSTM) are particularly well-suited for analyzing time-series data, such as ECG signals, owing to their capability to capture sequential information. LSTMs, with their capacity to retain long-term dependencies, are extensively utilized to predict the occurrence of arrhythmias or other cardiovascular conditions by effectively identifying temporal patterns within the ECG data.

#### 4.3. Hybrid Deep Learning Approaches

Hybrid models, such as the Stacking Ensemble Learning Technique, combine traditional ML and DL techniques. They offer a promising approach to improving prediction accuracy.

##### 4.3.1. Stacking Ensemble Learning Technique

The ANN model is designed for learning from preprocessed data and accurately forecasting the prospect of heart disease in patients, and its mathematical framework is inspired by the complex architecture of the human brain, which consists of neurons with components such as axons, dendrites, and synapses. The proposed system integrates a three-

layered ANN model, which consists of an input layer, a hidden layer, and an output layer (Shakkeera et al., 2024). This model has been applied as a web application utilizing the Flask web framework, which is accessible to any user, medical provider, or patient for the purpose of predicting the likelihood of heart disease. Fig. 3 shows the integration of ANNs into the Stacking Ensemble Learning Technique, which enhances its performance in predicting heart diseases. ANNs, known for their capacity to capture complex nonlinear relationships in data, serve as powerful additions to ensemble-based models. During the ensemble training phase, alongside baseline classifiers such as Decision Forest, SVM, Random Forests, and probabilistic classifier. Based on their input features, trained ANNs use their algorithms to generate predictions. These predictions, along with those from other base models, are fed into the meta-model for training purposes. The meta-model also uses these inputs to produce a final ensemble output, whether it be a classification decision or a probability score for the prospect of heart disease. This integration capitalizes on ANN's ability of ANNs to discern intricate data patterns, thus complementing the strengths of simpler models. Through hyper-parameter tuning and optimization techniques, ANNs within the ensemble framework further refine predictive accuracy and robustness. Thus, by leveraging ANNs in Stacking Ensemble Learning, healthcare practitioners can gain a potent tool for precise heart disease prognosis, enabling proactive interventions and improving patient outcomes.

#### 4.4. Detection of Cardiovascular Disease during The Embryonic Stage

Cardiovascular disease is a fundamental aspect of preventive healthcare in early detection, and the Stacking Ensemble Learning Technique, in combination with ANNs, reinforces this objective. Fig. 4 shows the CVD prediction analysis and highlights the primary objective of detecting CVD using a dataset with missing values to pinpoint critical features, such as age, hypertension, sex, and hyperlipidemia. After identifying these features, preprocessing techniques were employed to address the missing values through imputation and feature scaling to ensure scale consistency. The method of finding pertinent data is known as feature extraction, such as demographic data (age and sex), clinical measurements (blood pressure and cholesterol), and

medical history (smoking status and family history), to build predictive models. Feature engineering refines this information by standardizing it, creating one-hot encodings for categorical variables, and generating interaction terms to capture complex relationships. Feature selection is guided by domain expertise, which ensures that the chosen features are both relevant and informative. Additional techniques, such as normalization, polynomial features, and dimensionality reduction, and also used to enhance the predictive authority of the dataset. The ultimate goal is to construct robust predictive models that can accurately diagnose disease for heart and aid clinicians in making decisions. Subsequently, various classifiers, including logistic regression and decision trees (Sharmila et al., 2017), were trained on their

preprocessed data to approximate the likelihood of heart disease. Evaluating these classifiers using accuracy, precision, and recall metrics enables identification of the top-performing model. Finally, the most effective classifier was implemented to predict whether an individual was likely to have heart disease based on their feature values, which ultimately leads to enhance patient outcomes and timely interventions via a comprehensive approach.

#### 4.5. Prediction Accuracy

Machine learning and deep learning approaches have emerged as critical tools for analyzing intricate health data. Combining ensemble learning techniques with SVM, RF, and ANNs aims to enhance the

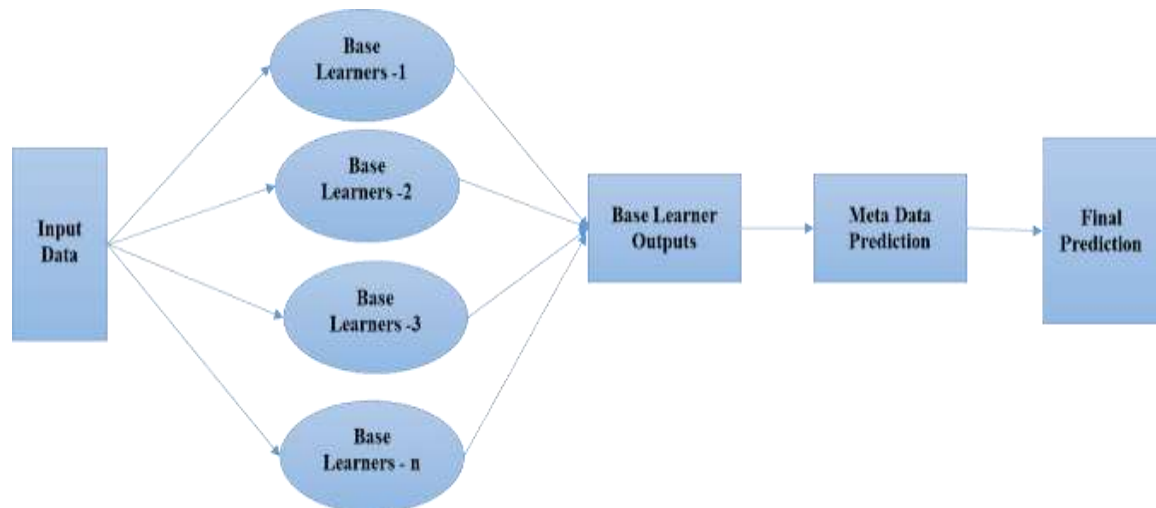


Fig. 3. Stacking Ensemble Learning Technique.

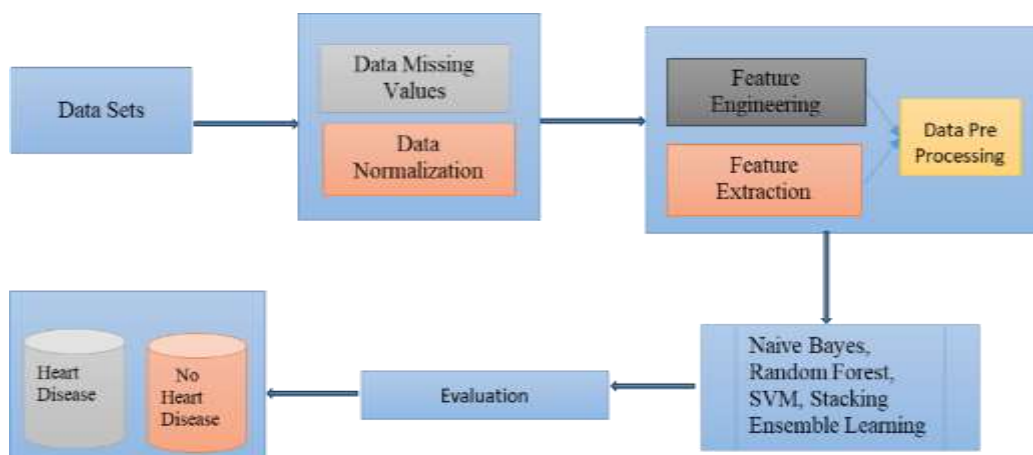


Fig. 4. Detecting cardiovascular disease.

predictive accuracy of heart disease detection by identifying subtle patterns in ECG data that may be overlooked by traditional analytical methods.

#### 4.6. Complexity and Noise

Preprocessing techniques, such as feature extraction, dimensionality reduction, and noise filtering, play a vigorous role in addressing the encounters posed by noisy and high-dimensional medical data, like ECG data, for ML models. These preprocessing steps help to identify the most relevant and informative features, reduce the complication of the input space, and mitigate the effects of noise and artifacts, ultimately improving the predictive performance of the models for heart disease detection.

#### 4.7. Early Detection and Intervention

The early identification of cardiovascular disorders is essential for mitigating severe health issues and enhancing patient prognoses. Advanced methodologies that leverage DL techniques aspire to analyze ECG data and other biometric indicators to

detect early signs and markers of CVD. By harnessing the capabilities of these sophisticated ML approaches, clinicians can acquire valuable insights that facilitate timely medical intervention and proactive management of heart-related conditions, ultimately improving the prospects of successful treatment and disease prevention.

#### 4.8. ANN-Based Prediction Advantages

Fig. 5 shows one of its key strengths in the model, which is its scalability. This capability enables the analysis of extensive datasets and population-wide evaluation (Hlaudi et al., 2014). It comprises 14 features, including age, gender, hypertension, tobacco consumption, hyperglycemia, abnormal ECG readings, and chest pain. This allows the identification of complex CVD risk factors and supports the implementation of targeted interventions on a larger scale. Another benefit of the ANN-based prediction model in an ensemble is its potential to reduce costs. By accurately identifying individuals at risk before disease progression to advanced stages, the healthcare expenses associated with severe treatments are

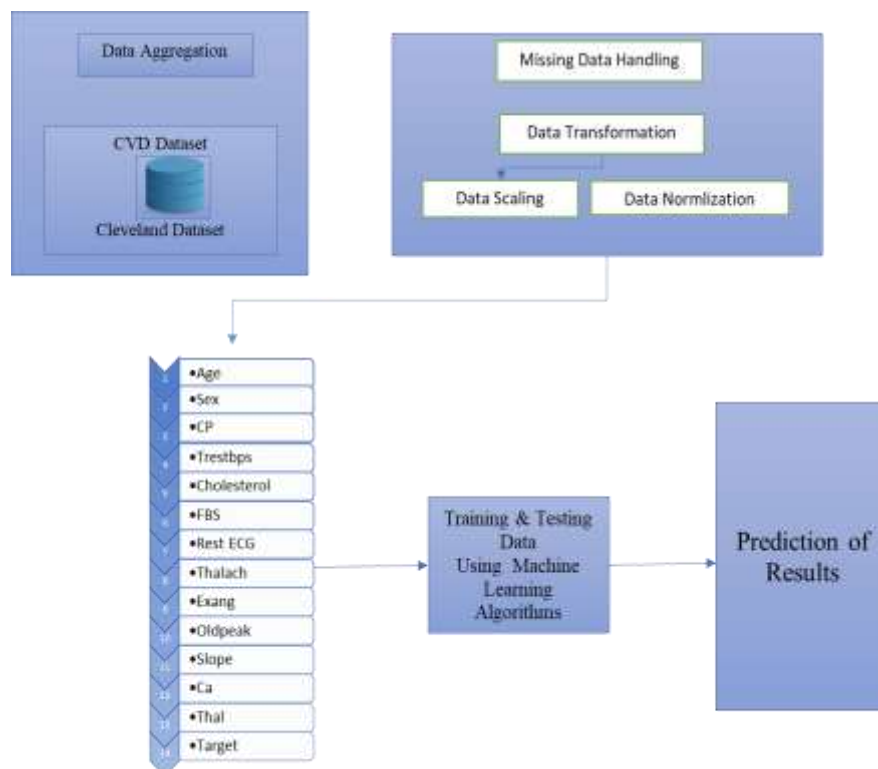


Fig. 5. Advantages of Artificial Neural Network.

minimized. This alleviates the financial burden on healthcare systems and financially benefits patients (Theresa, 2016).

Ultimately, the ANN-based prediction model within the ensemble aims to enhance patient outcomes. By accurately identifying at-risk individuals and facilitating early interventions, the ensemble reduces mortality rates and significantly improves patient well-being (Animesh et al., 2017). This transformative approach to cardiovascular healthcare highlights the pivotal role of the ensemble in improving patient outcomes and the overall healthcare system.

Fig. 6 shows that the proposed system has an impressive accuracy rate of 91.1%, thereby surpassing other existing models. It is also worth noting that the proposed system model may be pre-trained on ImageNet.

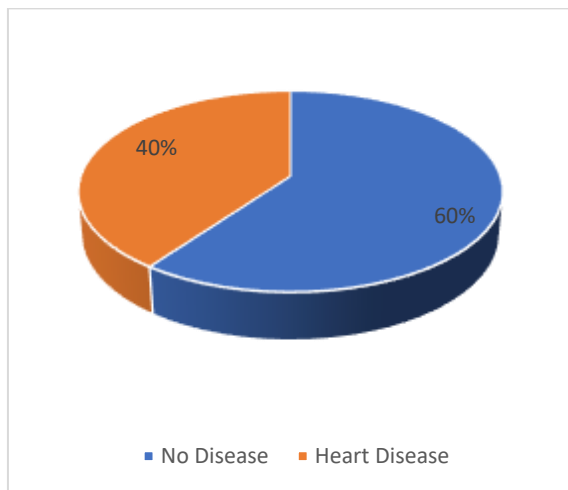


Fig. 6. Cardiovascular disease prediction.

## 5. Results and Discussion

This module evaluates the ANN model results and predicts the likelihood of heart disease in patients by using an actual dataset. Table 1 describes 14 features, including age, gender, hypertension, tobacco consumption, hyperglycemia, abnormal ECG readings, FBS, thalach, Exang, Oldpeak, Ca, Thai, and Target. Fig. 7 shows the result of training and validating the model once completed, and its performance was assessed using the test set, with the test set accuracy serving as a crucial performance metric.

### 5.1. Model Evaluation Using Confusion Matrices

In this study, performance evaluation metrics are confusion matrices generated for each ML algorithm employed, including SVM, Random Forests, ANNs, and hybrid models like the Stacking Ensemble Learning approach. Depth analysis provided a nuanced understanding of how effectively each model could distinguish between patients with and without heart disease.

### 5.2. Accuracy

Accuracy characterizes the total percentage of correct predictions made by the model. Despite a crucial metric, it can be misleading when dealing with imbalanced datasets, such as those commonly found in healthcare, where the number of non-disease cases may outweigh the disease cases.

### 5.3. Precision (Positive Predictive Value)

Precision represents the fraction of positive predictions that are truly correct. In medical applications, precision is crucial to minimize false alarms, ensuring that only patients at genuine risk are identified.

### 5.4. Recall (Sensitivity)

Recall measures the model's capability to correctly identify actual positive instances. A high recall value indicates that the model is effective at detecting the majority of individuals with heart disease.

### 5.5. F1-score

The F1-score represents the harmonic mean of the model's precision and recall. This metric is particularly valuable when seeking to strike a balance between precision and recall, especially in scenarios where the class distribution within the dataset is skewed or imbalanced. Fig. 8 shows a performance analysis of the confusion matrix, and Fig. 9 shows an overall comparison of the accuracy of various models. Our model Stacking Ensemble Technique achieved 99%, which involves the trained ANN model for future utilization, such as deploying it as a web application that healthcare professionals and patients can access to foresee the probability of heart diseases.

**Table 1.** Detailed description of the dataset, including its value ranges and data types.

No.	Feature Name	Feature Code	Description	Value Ranges	Data Type
1	Age	Age	Age in years completed	between 29 and 77	Numeric
2	Sex	Sex	Male: 1, female: 0	0 or 1	Nominal
3	Type of chest pain	CP	Typical angina: 1, atypical angina: 2, non-angina pain: 3, asymptomatic: 4	1 to 4	Nominal
4	Resting blood pressure	Trestbps	Patient's Resting Blood Pressure Range	94 to 200 mm Hg	Numeric
5	Serum cholesterol	Chol	Cholesterol level in mg/dl	126 to 564 mg/dl	Numeric
6	Fasting blood sugar	Fbs	Fasting Blood Sugar >120 mg/dl (true:1, false: 0)	0 or 1	Nominal
7	Resting electro-cardiographic results	Restecg	Normal: 0, ST-T wave abnormality:1, Hypertrophy: 2)	0, 1 and 2	Nominal
8	Maximum heart rate	Thalach	Heart Rate of Patients	71 to 202	Numeric
9	Exercise-induced angina	Exang	Patient experienced angina during exercise(Yes=1, No=0)	0 or 1	Nominal
10	ST depression induced by exercise relative to rest	Oldpeak	Depression caused by exercise, Up sloping: 1, Flat: 2, down sloping: 3	1 to 3	Numeric
11	The slope of the peak exercise	ST segment Slope	Slope of peak exercise	1, 2, 3	Nominal
12	Number of major vessels (0-3) colored by fluoroscopy	Ca	Major Vessels colored by fluoroscopy with range 0 to 3	0 to 3	Numeric
13	Thallium	Thal	Represents thallium stress test, Normal:3, fixed defect: 6, reversible defect: 7	3, 6, 7	Nominal
14	Target	Target	Output, Heart disease present: 1, heart disease absent: 0	0 or 1	Nominal

## 6. Conclusions and Future Scope

In conclusion, predictive analytics, powered by ML and DL, holds immense promise in unlocking the hidden value within large, complex datasets. As the field of predictive analytics continues to evolve, addressing the challenges of uncertainty, overfitting, and accessibility will be crucial to realizing its full potential.

By continually improving the reliability, transparency, and scalability of predictive analytics techniques, organizations can harness the power of data-driven insights to drive strategic decision-making, enhance operational efficiency, and, ultimately, improve outcomes across a wide range of domains. In future research, incorporating wearable sensor technologies for continuous real-time monitoring of cardiovascular health is a promising area of study. Also, discovering progressive DL



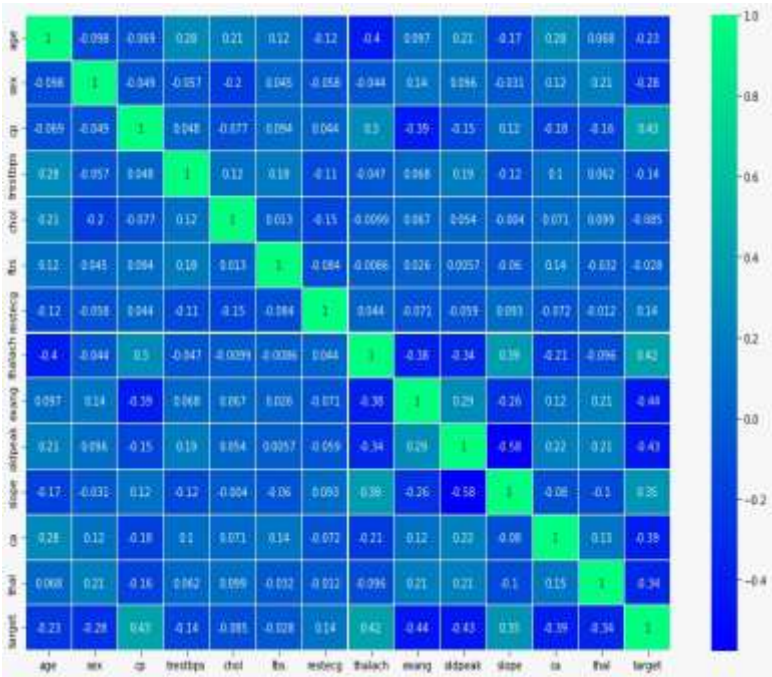


Fig. 7. Exploring correlated features in cardiovascular disease.

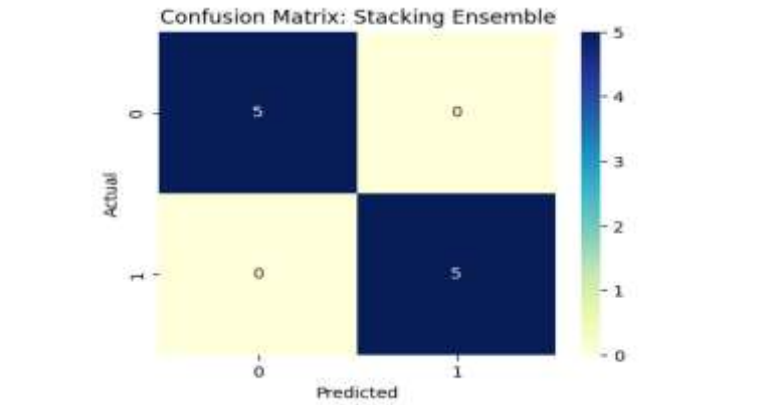


Fig. 8. Performance analysis of confusion matrix.

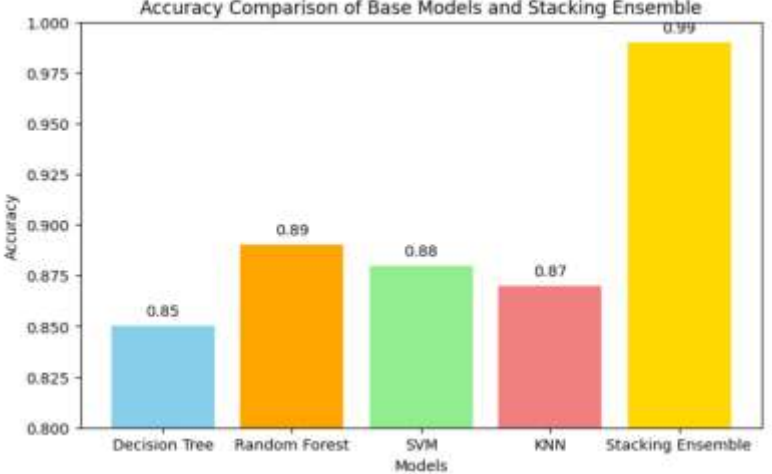


Fig. 9. Accuracy of stacking ensemble technique.



methods, such as graph neural networks for analyzing ECG signals, could further enhance the accuracy of disease prediction. Expanding the dataset to include diverse populations would improve the generalizability of the models. Investigating the potential of federated learning to enable privacy-preserving cardiovascular predictions is another valuable research direction. Finally, personalized healthcare interventions driven by AI insights could revolutionize early diagnosis and treatment of cardiovascular conditions. The Cleveland Cardiovascular Disease database was employed for training and estimating the algorithm, which resulted in the development effective model among various imperative-based combinations. Furthermore, we conducted a comparative analysis of accuracy and identified the optimal and balanced model.

## References

- Agrawal S., Revolutionizing Cardiovascular Health: A Machine Learning Approach for Predictive Analysis and Personalized Intervention in Heart Disease, *Int J Res Appl Sci Eng Technol*, vol. 12, no. 3, pp. 233–246, Mar. 2024, doi: 10.22214/ijraset.2024.58797.
- Animesh Hazra, Subrata Kumar Mandal, Amit Gupta, Arkomita Mukherjee & Asmita Mukherjee, Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review, *Advances in Computational Sciences and Technology*, ISSN 0973-6107 Volume 10, Number 7 (2017) pp. 2137-2159
- Bellman, R.E. & Zadeh, L.A. (1970). Decision-making in a fuzzy environment, *Management Science*, 17(4), 141-164.
- Chaudhary I. & Anwar H., The Impact of Artificial Intelligence on Cardiovascular Disease Diagnosis: A Review, *Pakistan Journal of Medical and Health Sciences*, vol. 17, no. 11, pp. 8–13, Feb. 2024, doi: 10.53350/pjmhs0202317118.
- Dangare C. S. & Apte S. S., Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques, *Int J Comput Appl*, vol. 47, no. 10, pp. 44–48, Jun. 2012, doi: 10.5120/7228-0076.
- Faldu P., Menaria S., Patel T., Mt.-C. Student, and A. Professor, “Issue 12 [www.jetir.org](http://www.jetir.org) (ISSN-2349-5162),” 2019. [Online]. Available: [www.jetir.org](http://www.jetir.org)
- Hlaudi Daniel Masethe et al., Prediction of Heart Disease using Classification Algorithms, *Proceedings of the World Congress on Engineering and Computer Science [WCECS]*, Volume: II, ISBN: 978-988-37253-7-4 ISSN: 2078-0958 (Print), ISSN: 2078-0966 (Online) (2014).
- Hsu, W.L. (2008). Decision Model of Product Design, Ph.D. Dissertation of Folks University, Taiwan. (in Chinese).
- Maganathan T., Senthilkumar S., & Balakrishnan V., Machine Learning and Data Analytics for Environmental Science: A Review, Prospects and Challenges, in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Nov. 2020. doi: 10.1088/1757-899X/955/1/012107.
- Mandal S., “Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review,” 2017. [Online]. Available: <http://www.ripublication.com>.
- Naseer A., Muhammad, Khan M., Arif F., Iqbal W., and Ahmad A., “An Improved Hybrid Model for Cardiovascular Disease Detection Using Machine Learning in IoT.
- Polaraju K., Durga Prasad D., and Tech Scholar M., Prediction of Heart Disease using Multiple Linear Regression Model, 2017. [Online]. Available: [www.ijedr.org](http://www.ijedr.org).
- Shakkeera L. & Kavitha P. (2024). Predictive Modelling for Medical Image Analysis Using Deep Learning Techniques. *International Journal on Recent and Innovation Trends in Computing and Communication*, 12(1), 291–298. from <https://ijritcc.org/index.php/ijritcc/article/view/10327>.
- Sharmila S. et al., Analysis of Heart Disease Prediction using Data Mining Techniques, *International Journal of Advanced Networking*

& Applications (IJANA), Volume: 08, Issue: 05,  
Pages: 93-95 (2017).

Srinivasa Rao G. & Muneeswari G., A Review:  
Machine Learning and Data Mining Approaches  
for Cardiovascular Disease Diagnosis and  
Prediction, EAI Endorsed Trans Pervasive Health  
Technol, vol. 10, Mar. 2024, doi:  
10.4108/eetpht.10.5411.

Theresa Princy R, Human Heart Disease Prediction  
System using Data Mining Techniques,  
International Conference on Circuit, Power and  
Computing Technologies [ICCPCT], IEEE (2016).

Thansekhar M. R. & Balaji N., Heart Disease  
Diagnosis using Predictive DataMining. [Online].  
Available:  
<http://www.cs.waikato.ac.nz/ml/weka>.

## AUTHOR BIOGRAPHIES



**Kavitha P** received her  
**M.E** degree from Apollo  
Priyadharsam Institute of Science  
and Technology, Chennai, India in  
2022. She is currently pursuing  
her Ph.D. in Computer Science  
and Engineering from Presidency  
University, Bengaluru, India. Her areas of research  
interest include Machine Learning, Deep Learning.



**Shakkeera L** received her Ph.D.  
degree from B.S. Abdur Rahman  
Crescent Institute of Science and  
Technology, Anna University,  
Chennai, India in 2018. She is  
currently a Professor & Associate Dean in the School  
of Computer Science and Engineering & Information  
Science from Presidency University, Bengaluru,  
India. Her areas of research interest include Mobile  
Cloud Computing, Machine Learning, IoT, MANET,  
Information Security, and Data Analytics.

# MNETGIDD: A heuristic-oriented segmentation and deep learning multi-disease detection model for gastrointestinal tracts

A. Bamini<sup>1\*</sup>

<sup>1</sup>Department of Computer Applications, The Standard Fireworks Rajaratnam College for Women, Sivakasi, Tamil Nadu, India

\*Corresponding author E-mail: drbaminia@gmail.com

(Received 6 May 2024; Final version received 18 September 2024; Accepted 2 January 2025)

## Abstract

Malignant growth of the gastrointestinal (GI) tract is among the leading causes of death worldwide. Research indicates that almost 40% of people worldwide suffer from long-term digestive issues. According to a study published in the United European Gastroenterology Journal, digestive disorders have increased since 2000. Digestive disorders continue to be a major cause of death, even with a slight decline. The World Health Organization's Mortality Database reported huge death rates every year due to GI diseases. From that report, the need to accurately detect GI tract malignant in low-cost and error-prone labor must be developed. This work introduces MNET Gastrointestinal Disease Detection (MNETGIDD), which is a complete identification model for multi-gastrointestinal disease discovery from clinical images. MNETGIDD model uses the Gastrolab dataset with endoscopic images, acting as pipelines that are pre-processed and segmented to identify the affected areas. This proposed approach aims to enhance image quality and facilitate accurate segmentation and classification through a pipeline process, initially preprocessing with techniques such as text removal, illumination enhancement, and fuzzy histogram equalization. During segmentation, Otsu segmentation based on Krill-Herd optimization was used to identify the affected area. The MNETGIDD model incorporates the MobileNetV2 architecture, designed for a lightweight classification model working under resource-constrained environments. According to the tests, the MNETGIDD model exhibits high sensitivity and specificity, often outperforming human experts. In terms of accuracy, the model achieved 96.349%, a precision of 96.25 %, and a recall of 97.08%. This deep learning system has the potential to revolutionize gastrointestinal disease diagnostics and screening by automating key steps and improving patient outcomes.

**Keywords:** Fuzzy Histogram Equalization, Gastrointestinal Disease Detection, MNETGIDD, Gastrolab dataset, Low-light Image Enhancement, Mean-Shift Segmentation, MobileNetV2

## 1. Introduction

In the world, cancer is the leading cause of death, and gastrointestinal (GI) cancer is the most common type. Globally, 1.8 million people die from gastrointestinal diseases each year (Sharmila & Geetha, 2022), and GI cancer is the fourth leading cause of death. Globally, gastrointestinal diseases are

a significant health burden. The Global Burden of Disease Study 2019 reports that digestive diseases cause over 2.5 million deaths worldwide, accounting for 4.5% of all deaths (Theo, 2019). Globally, digestive diseases caused 81.1 million disability-adjusted life years (DALYs).

The growth of GI polyps on the mucosa of the stomach and colon is the cause of gastrointestinal cancer. The esophagus, stomach, small intestine, large intestine, rectum, and anus are some of the parts of the digestive tract affected by multi-GI diseases. A chronic disease can significantly impact an individual's quality of life. The symptoms, causes, and treatments of multi-gastrointestinal diseases differ from one type to the next. An estimated 1.9 million new cases of colorectal cancer, one of the four diseases examined in this study, will be diagnosed in 2020 (Sung et al., 2020). Our research also focused on gastric cancer, which was responsible for over 768,000 deaths worldwide in 2020. There were 600,000 cases of esophageal cancer worldwide in 2020. Approximately 6.8 million people worldwide suffer from inflammatory bowel diseases, including Crohn's disease and ulcerative colitis (Alatab et al., 2020). GI diseases have a substantial economic impact. According to a 2015 study, digestive diseases cost the United States \$136 billion annually in direct and indirect costs (Perry et al., 2020). Many diseases and conditions can affect the esophagus. Inflammation of the esophageal lining can cause heartburn and swallowing difficulties. Another common condition is hemorrhoids, which are caused by inflammation of the blood vessels in the anus and rectum. In different parts of the digestive tract, polyps can develop, and while most are benign, some are cancerous. A type of inflammatory bowel disease, ulcerative colitis grade 1 causes inflammation and ulcers in the colon and rectum linings. Because of their complexity, these disorders can be challenging to diagnose and treat; imaging, lab testing, and invasive diagnostic procedures might be required. Recently, a promising approach for diagnosing and segmenting multiple gastrointestinal diseases was developed using deep learning techniques. A computer-aided automated approach may be useful for highly accurate polyp diagnosis and cancer detection. Artificial intelligence (AI) holds immense promise in helping people visualize diseases that are invisible to the human eye in various medical fields (Ekiri et al., 2016). Endoscopy images can be evaluated, and key features of micro-imaged structures can be identified using AI tools. The following have been made possible by this research.

- The proposed scheme identifies and classifies the different GI diseases.

- By combining segmentation using IP and identification techniques via deep learning, a novel technique was proposed.
- Using this method, images for training and testing are randomly selected in the segmented images without manual intervention

### 1.1. Aim & Motivation

Medical practice and healthcare systems worldwide can be improved through computer-assisted early disease detection. Multi-GI disease segmentation and identification is a challenging but important task that could improve healthcare outcomes for millions of people worldwide. A deep learning model can enable us to accurately identify and segment complex diseases like esophagitis grade A, hemorrhoids, polyps, and ulcerative colitis grade 1.

Improved patient outcomes depend on early detection and accurate diagnosis of gastrointestinal diseases. In a study by Smith et al. (2022), early detection of colorectal cancer can increase 5-year survival from 14% to 90%. This study proposed the MNET Gastrointestinal Disease Detection (MNETGIDD) model as an advanced diagnostic tool. A recent advancement in AI and deep learning has shown promising results. Based on a meta-analysis conducted by Johnson et al. in 2023, AI-assisted diagnosis of GI diseases could improve detection rates by 30%.

This work aimed to develop a framework, MNETGIDD, for recognizing a wide range of GI diseases simultaneously rather than multiple tools used individually to detect an anonymous disease. A significant improvement over most recent studies is that this study considers different types of GI diseases (both diseased and normal cases) related to the human GI tract. Furthermore, the model proposes a means of identifying diagnostic decisions based on deep learning techniques. A medical expert can validate the computer decision interactively with the assistance of such additional information. To address this problem, this work proposed a high-performance classifier and retrieval framework that uses endoscopic images to determine GI diseases using recent AI techniques.

### 1.2. Objective, Challenges and Issues

Deep learning techniques will be used to develop a model capable of accurately identifying the four common multi-gastrointestinal diseases: esophagitis

grade A, hemorrhoids, polyps, and ulcerative colitis grade 1. Deep learning will be used in conjunction with imaging data to train the proposed model to recognize patterns that may indicate the presence of these diseases. With the development of this model, the need for invasive diagnostic procedures will be reduced, and patient outcomes will improve. A multi-gastrointestinal disease model can also help healthcare professionals diagnose and manage these illnesses. A computer-aided diagnosis system was developed to aid medical experts in diagnosing different types of gastrointestinal diseases.

A model for segmenting and identifying multi-GI diseases is a challenging task that requires consideration of several ethical, practical, and technical factors. There are several challenges and issues associated with this research, including:

- **Limited availability of labeled medical imaging data:** Medical imaging data labeled with medical conditions is necessary to develop a reliable and accurate model for segmenting and identifying multi-GI diseases. Multi-GI diseases can be segmented and identified inaccurately due to variations in medical imaging data regarding quality, resolution, and imaging modalities. Imaging data is often variable, making it difficult to generalize models to new and unknown data.
- **Technical complexity of deep learning techniques:** Developing and implementing deep learning techniques like neural networks and fuzzy logic require high levels of technical expertise. Developing these techniques can be difficult in resource-constrained environments due to the high computational requirements.
- **Ethical considerations:** Medical imaging data can raise ethical concerns like confidentiality and patient privacy. The protection of patient data requires compliance with ethical guidelines and regulations. Medical imaging data must be stored securely and accessed only by authorized personnel. The data should be used only for the stated purpose and must not be shared without the patient's consent. All patient data must be deleted when no longer needed.
- **Integration with clinical workflows:** A successful clinical practice is one that integrates the proposed model seamlessly into existing clinical workflows. The research must examine practical and operational aspects for the model to

be usable and acceptable to healthcare professionals.

### 1.3. Contribution of the Work

The proposed MNETGIDD mode contains three phases: pre-processing, segmentation, and classification, which help identify the affected area from different types of GI tracts. Each phase's contribution is explained as follows:

- To improve the image quality during pre-processing using the in-painting, low-light image enhancement (LIME), and fuzzy logic-based histogram equalization (FHE) algorithms.
- To apply the KHO-Otsu automated threshold-based algorithm to segment the affected area in different disease images.
- The MNET Model is utilized to identify the impacted images within the test set.

### 1.4. Structure of the Paper

Section 2 reviews the relevant literature, including GI-based pre-processing techniques and ideas, segmentation techniques, machine learning, and deep learning approaches. Section 3 describes the proposed scheme of this work and how data was collected, pre-processed, segmented, and classification modeling techniques were used. The results and discussion for pre-processing, segmentation, and classification techniques employed for the input dataset with its performance evaluation are explained in Section 4. Finally, the conclusion of this work is described in Section 5.

## 2. Review of Literature

Sharmila & Geetha (2022) proposed a deep learning model that combines a deep CNN with a pre-trained model, ResNet101, to detect and classify abnormalities in the GI tract. A goal of the proposed research is the detection of disease in endoscopic images. A public dataset of 8,000 images called KVASIR forms the basis of the architecture. An accuracy of 98.37% was achieved using the convoluted neural network (CNN) approach. A higher level of recognition is achieved without an individual's assistance in the experiment.

The dataset used by An et al. (2022) contained data on eight diseases. The data set and parameters used in the literature were also compared with those used in other studies. The study's results section also provides detailed classification results for eight diseases. Wong et al. (2022) implemented deep transfer learning for classifying GI diseases. Cleansing, standardizing, and transforming data are performed after exploratory analysis. To solve the classification problem, a pre-trained ResNet50 is utilized, which is a CNN with 50 layers. Benchmarking and performance metrics were used to evaluate the model. According to the results, the proposed model showed high stability with consistent scores.

According to Nguyen et al. (2022), upper GI tract diseases can be automatically classified using a computer algorithm. There are two main components to the framework: a CNN based on the ResNet-50 architecture and a Focal Loss application, as well as a geometric transformation, brightness and contrast transformation, for revealing hidden characteristics of upper GI diseases and anatomical landmarks and for dealing with imbalanced datasets.

Su et al. (2022) demonstrated that their proposed method outperforms existing computational methods on GI disease screening benchmark datasets. Sharib et al. (2021) developed multiple methods to tackle two sub-challenges in clinical endoscopy: artifact detection and segmentation and disease detection and segmentation. Datasets from clinical endoscopy were used and algorithms were evaluated for generalization ability. Despite most teams focusing on accuracy, only a few clinical use methods are considered credible. Ekiri et al. (2016) evaluated a real-time polymerase chain reaction (PCR) assay for detecting *Salmonella* in fecal samples from hospitalized horses with and without GI disease symptoms. *Salmonella* in the feces of horses can be detected with PCR assays targeting the *Salmonella invA* gene. The possibility of nosocomial *Salmonella* infections can also be detected through further bacteriologic culture testing. Several GI diseases have been classified, segmented, and detected by automated methods in a Naz et al. study (2021). An in-depth description of these state-of-the-art methods is presented in the paper. Moreover, literature is categorized according to the method used for preprocessing, segmentation, and handcrafted features-based methods.

A computer-aided detection method for lower gastrointestinal diseases was developed by Al-

Adhaileh et al. (2021) using modified criteria for extracting deep shape, color, and texture features and adapting them to a transfer method for fine-tuning and contouring. Extensive experiments were conducted to diagnose lower gastrointestinal diseases. A new model was developed for transferring features from a nonmedical deep learning dataset and adapting them to a medical dataset. Yogapriya et al. (2021) integrated traditional image processing algorithms and data augmentation techniques with a fine-tuned pre-trained deep CNN to classify GI diseases using images captured by wireless endoscopy. Concatenating VGGNet and InceptionNet networks for the purpose of developing a model to diagnose gastrointestinal diseases was proposed by Melaku et al. (2022). Deep convolutional neural networks VGGNet and InceptionNet are trained and used to extract features from endoscopic images. By concatenating these extracted features, machine learning (Softmax, k-nearest neighbor, random forest, and support vector machine [SVM]) classification techniques were used to classify them. Based on the available standard dataset, SVM was found to perform better than the other techniques.

An endoscopy image classification system based on CNN was proposed by Ramamurthy et al. (2022). Effimix is a CNN architecture that combines state-of-the-art technology (such as EfficientNet B0) with custom-built architectures. The proposed Effimix model employs squeeze and excitation layers and self-normalizing activation layers to classify GI diseases accurately. The proposed architecture has been tested on the HyperKvasir dataset for the classification of endoscopy images. A spatial factor can be used to improve the performance of classification, according to Lonseko et al. (2021). In particular, the proposed mechanism uses encoder-decoder layers to implement a CNN-based spatial attention mechanism for classifying GI diseases. Our data-augmentation techniques help solve the problem of data imbalance. This method was validated using 12,147 multi-sited, multi-diseased GI images from publicly available and private sources.

An approach based on ResNet-50 architecture was proposed by Gammulle et al. (2020). A relational network was used to classify abnormalities in the human gastrointestinal tract using endoscopic images based on features extracted from a pre-trained mid-layer model. In a recent study (Cogan et al., 2019), the authors employed NASNet, Inception-v4, and Inception-ResNet-v2 architectures to recognize

anatomical landmarks and diseased tissue. An approach is proposed using endoscopic images to remove the edges, filter, and enhance contrast, scale, and color map. Their MCC for identifying eight abnormalities of the human gastrointestinal tract was 0.93. An abnormality detection method was proposed by Jain et al. (2020). In the first phase of that work, they extracted useful features from images using fractal dimension techniques. They used random forest classifiers to classify abnormal endoscopic images. Jain et al. (2021) proposed an attention-based model to classify endoscopic images into four categories. Once the endoscopic image has been identified as abnormal, the second stage of that work is based on anomaly detection. Recent reviews of GI tract abnormalities with endoscopic images (Jha et al., 2021) indicated that the manual assessment of a large number of gastric images is a laborious job and needs expertise. Computer-aided diagnosis methods can be developed to handle the dilemma of manual analysis of the substantial volume of endoscopic data.

In a study by Gunasekaran et al (2023), they combined DenseNet201, InceptionV3, and ResNet50 to obtain 94.54%, 88.38%, and 90.58% accuracy, respectively. Model averaging and weighted averaging are used to combine predictions. A model-averaging ensemble has an accuracy of 92.96%, while a weighted average ensemble has an accuracy of 95.0%. The weighted average ensemble outperformed all models. As a result of the evaluation, they correctly classified features using an ensemble of base learners.

An objective comparison of state-of-the-art methods versus those developed by participants for two sub-challenges is provided by Sharib et al. (2021) for artifact detection and segmentation (EAD2020) and disease detection and segmentation (EDD2020). Data were collected for both EAD2020 and EDD2020 sub-challenges across multiple centers, organs, classes, and modes. Algorithms were also evaluated for out-of-sample generalization. Despite most teams focusing on accuracy, only a few methods have clinical validity. Exploring data augmentation, data fusion, and optimal class thresholding techniques, the top-performing teams tackled class imbalance and variabilities in size, origin, modality, and occurrences.

Sharma et al. (2023) propose that data augmentation strategies and statistical measures have been used to improve the model's performance. A total of 1,200 images were used in the test set to assess accuracy and robustness. A CNN model trained with ResNet50 weights averaged 99.80% accuracy on the

training set (100% precision and 99% recall) and 99.16% accuracy on the validation and additional test sets, respectively, while diagnosing GI diseases.

## 2.1. Limitations and Research Gap

1. Most existing methods focus on increasing accuracy, but few are applied in clinical settings. As Sharib et al. (2021) have pointed out, credible methods must be developed to be used reliably in clinical settings.

2. While most studies focus on a limited number of GI disorders, there may be a wider range of disorders and diseases in clinical settings that should be identified and classified correctly.

3. Many proposed methods use complex deep learning architectures and require significant computational resources, which are not always readily available in clinical settings. More efficient and lightweight models are needed for resource-constrained devices.

4. Manually assessing many endoscopic images requires expertise, as Jha et al. (2021) noted. The development of reliable computer-aided diagnosis methods that can reduce this burden is an important research gap.

## 3. Proposed Model

The complex nature of multi-gastrointestinal diseases can make identifying and segmenting them difficult. These diseases can, however, be segmented and identified using image processing and deep learning techniques. Fig. 1 shows the architecture of the proposed **MNETGIDD** model.

There are a number of conditions affecting the upper and lower digestive tracts which are included in this dataset. These conditions are included:

- Symptoms of esophagitis include throat and stomach inflammation.
- Pain, itching, and bleeding are common symptoms of swollen veins in the lower rectum and anus.
- A growth that develops in the lining of the colon or rectum that is abnormal in nature.
- Symptoms of ulcerative colitis include colon and rectal inflammation and ulcers.

The Gastrolab dataset can be used to assess the prevalence, symptoms, risk factors, and affected



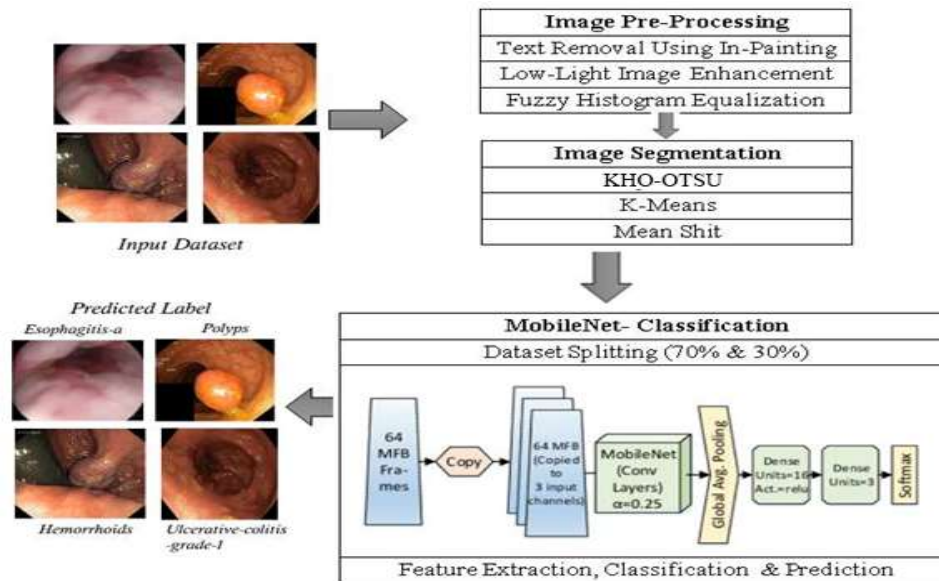






Fig. 1. Proposed work of MNETGIDD architecture.

Table 1. Sample image set for different gastrointestinal tract.

Esophagitis grade A	Hemorrhoids	Polyps	Ulcerative colitis grade 1
			
Upper gastrointestinal tract	Lower gastrointestinal tract	Lower gastrointestinal tract	Lower gastrointestinal tract

datasets related to these common gastrointestinal conditions. Gastrolab provided a collection of endoscopic videos, which included endoscopic videos of both normal and diseased GI tracts. The names of the videos include information about the normal and diseased cases, as well as the anatomical district. The dataset was classified into 37 different classes based on the available information. Among the 37 different classes, this work extracts four serious problems of diseases. An image dataset is created from the video. Using this data, a predictive model or decision support tool related to gastrointestinal health could also be developed. Table 1 shows the sample images from the dataset.

### 3.1. Image Preprocessing

Preprocessing the images enhances low light, removes the text, and improves the quality as the first step in the model. Improving image quality and preparing images for further analysis, such as segmentation, is necessary. this pre-processing phase involves removing text from an image, enhancing the

image with low-light illumination, and applying fuzzy histogram equalization.

#### 3.1.1. Text Removal Using In-Painting

The image is preprocessed by removing any text or annotations that may be present. The presence of text may complicate a text-based segmentation and classification task, as the model may learn features related to the text rather than the actual medical pathologies. A connected component analysis identifies discrete regions within the image after the regions have been identified. Thus, Text regions are distinguished from the rest of the image content. Connected components are analyzed to identify regions likely to contain text. It is possible to differentiate text regions from other image features based on their aspect ratio, size, and intensity contrast. A technique called painting is used to remove the text regions from the image once they have been identified. In painting, the identified text areas are eliminated while preserving the underlying medical content by filling it with surrounding pixel values.

**Table 2.** Text removal



This work used a connected component analysis with a minimum area threshold of 50 pixels for text removal to identify potential text regions. The aspect ratio for text components was set between 0.2 and 5.0. We applied an exemplar-based in-painting algorithm with a patch size of 9x9 pixels and a search window of 81x81 pixels to fill in the removed text areas

### 3.1.2. Low-light Image Enhancement

Low-light is common for images captured in the GI tract at low-light conditions to have a low level of

visibility. Images captured in low light conditions are barely satisfactory, for one thing. Low-light images may be included in the dataset acquired for disease identification of gastrointestinal images. When the camera takes photographs with light entering the human body's intestinal system, which is often in poor light conditions, the images can be degraded. This not only affects the recognition but also the performance of computer-based applications. A low-light image can be enhanced by estimating its illumination map through a method called "Low-Light Enhancement through Illumination Map." An illumination map is used to estimate the amount of illumination in different regions of an image, and then this map is used to adjust brightness and contrast. In Max-RGB, the highest value is identified between the three-color channels (R, G, and B) in an attempt to estimate illumination. Its effectiveness is limited to enhancing global illumination and enhancing global illumination only.

**Algorithm 1.** Text removal process.

#### Algorithm Text Removal

**Input:**  $Img_{In}$  – Input Image

**Output:**  $Img_{Out}$  – Output Image

1. Read the input Image  $Img_{In}$
2. Components ( $C_s$ ) = obtain connected regions ( $Img_{In}$ )
3. Text\_regions ( $T_R$ ) = [ ]
4. for component (C) in  $C_s$ :
5. if  $T_R$  (component) is Available:
6.  $T_R.append(component)$
7.  $temp = Img_{In}.copy()$
8. for Text\_region in  $T_R$ :
9.  $Img_{Out} = Inpaint(Img_{Out}, Text\_region)$
10. return  $Img_{Out}$

**Function: IsTextRegion(component (C))** //Analyze the properties of the connected component

1. aspect\_ratio ( $A_R$ ) =  $C.width / C.height$
2. size = component.area
3. contrast = Calculate Contrast(component)
4. if ( $A_R > 2$  and size < 1000 and contrast > 0.5:
5. return True
6. else:
7. return False

**End Algorithm**

**Algorithm 2.** Low-light image enhancement through illumination map.

```

Algorithm LIME
{
  Input:  $Img_{In}$  as Image Set
  Output:  $Img_{Out}$  as Output Image
  1:   Read the input Image  $Img_{In}$ 
  2:   Convert to grayscale,  $Img_{Gray} =$ 
  Grayscale( $Img_{In}$ )
  3:   Generate illumination map
  4:    $I_{Map} =$ 
  Generate_IlluminationMap( $Img_{Gray}$ )
  5:   Apply multi-scale
  decomposition to the illumination map
  6:    $D_{Map} =$ 
  Decompose_MS(illumination_map)
  7:   Generate
  Illumination_Adjustment_Maps  $A_{Map} = []$ 
  8:   For the map in  $D_{Map}$ 
  a.    $A_{Map} = \text{Normalize}(\text{map})$ 
  b.    $A_{Map} \cdot \text{Append}()$ 
  9:   Combine adjustment maps into a
  final map  $F_{Map}$ 
  10:   $Img_{En} =$ 
  Apply_Illumination_Adjust( $Img_{In}$ ,  $F_{Map}$ )
  11:  Convert back to original color
  space
  12:   $Img_{Out} =$ 
  convert_color_space( $Img_{En}$ ,
  original_color_space)
}

```

Using tone mapping, the illumination map can be applied to an image and adjusted to reduce the size of its dynamic range, allowing it to be displayed on devices with a low dynamic range. It is intended to improve low-light image quality and visibility through the use of illumination maps. By considering adjacent pixels within a small area surrounding a specific target pixel, the majority of these enhancements focus on local illumination uniformity.

This work estimated the illumination map using a multi-scale Retinex algorithm in the LIME process. It used three scales (15, 80, and 250) with respective weights of 0.3, 0.5, and 0.2. The final illumination map was refined using a guided filter with a radius of 15 pixels and a regularization parameter of 0.001. We applied gamma correction with  $\gamma = 0.6$  to enhance the contrast of the illumination-adjusted image.

**Table 3.** Low-light image enhancement.



### 3.1.3. Fuzzy Histogram Equalization

Enhancing an image is typically aimed at revealing hidden details in the image or enhancing its contrast by extending its dynamic range. An image contrast enhancement technique commonly used is histogram equalization (HE). The HE algorithm involves remapping the gray levels of an image based on the probability distribution of its input gray levels. FHE addresses issues such as over-enhancement and noise amplification in grayscale and color images. A fuzzy histogram is divided into two portions based on its median value. Then HE approaches are applied to each sub-histogram independently to enhance local contrast while maintaining image brightness. Using membership functions, fuzzy images are mapped to a fuzzy plane, modified for contrast enhancement, and mapped back to gray levels using the fuzzy plane. A higher contrast image is generated by prioritizing gray levels that are closer to the original image's mean gray level. By using FHE, the brightness of the image is maintained and the local contrast is enhanced. To handle gray-level values more effectively, create a fuzzy histogram using fuzzy logic. A new dynamic range is assigned to each sub-histogram based on the median value of the original histogram. Each sub-histogram is then independently processed using the HE approach.

An image contrast can be enhanced while brightness is maintained using FHE. Each pixel intensity value is represented by a fuzzy membership function, which is used to compute the fuzzy histogram, divide it into two sub-histograms according to the median value, equalize each sub-histogram separately, and combine them to create a final fuzzy equalized histogram. The technique minimizes noise amplification and over-enhancement.

For the FHE, it defined a triangular membership function with parameters  $a = 0$ ,  $b = 128$ , and  $c = 255$  for the input gray levels. The fuzzy histogram was divided into two sub-histograms at the median value.

It used a modification factor  $\alpha = 0.5$  to control the degree of enhancement. The output membership function was defuzzified using the centroid method.

**Algorithm 3.** Fuzzy logic-based histogram equalization.

**Algorithm: FHE**  
 {  
**Input:**  $Img_{In}$  - Image Set  
**Output:**  $Img_{Out}$  - Output Image  
 1: Read the input Image  $Img_{In}$   
 2: Calculate the fuzzy membership (FM) function for each pixel intensity value using fuzzy logic  
 3:  $FM(x) = 1 / 1 + (\frac{|x-\mu|}{\sigma})^{2m}$   
 4: Compute the fuzzy histogram (FH) using the FM function for each pixel intensity (PI) value  
 5:  $FH(x) = \sum_{i=1}^N Membership(x_i).Frequency(x_i)$   
 6: Determine the median value of the original  $Img_{In}$  and divide the FH into 2 sub-histograms (SH) based on this value  
 7: Assign a new dynamic range to each (SH) and apply histogram equalization (HE) to each sub-histogram separately to enhance contrast  
 8: Combine the SH's to produce the final fuzzy equalized histogram  
 9:  $Img_{Out}(x) = InverseMembership(EqualizeHistogram)$   
 10: Convert the output image back to the original  
 11: Return  $Img_{Out}$   
 }

### 3.2. Image Segmentation

In the next step, a segmentation algorithm will be used to segment the image into regions of interest. Images can be segmented into multiple regions or segments corresponding to different objects or parts of them. A segmented image can be analyzed more easily and effectively by simplifying and/or transforming its representation. Computer vision, object recognition, scene analysis, and medical imaging are some of the applications that can benefit from this. Besides thresholding, edge detection, and clustering, there are a number of segmentation methods. The following methods are used to divide up

the affected area of the GI disease category. For instance, an image with a polyp's illness is portioned into impacted regions to recognize the illness.

#### 3.2.1. Mean-Shift Segmentation

In computer vision, images are segmented into several regions or segments with comparable attributes using mean shift segmentation. The mean shift algorithm works by finding the densest areas in the feature space of the image and classifying pixels based on how close each mode is to the other. This method's ability to handle nonlinear and nonparametric feature spaces gives it greater flexibility than some other segmentation methods. Among its uses in computer vision and image processing are object detection, tracking, and image compression.

**Algorithm 4.** Mean-shift segmentation.

**Algorithm MSS**  
 {  
**Input:**  $Img_{En}$  as Enhanced Image Set  
**Output:**  $Img_{Seg}$  as Segmented Image  
 1: Define the window size (W) and the kernel function (K)  
 2: Initialize all pixels (P) as unvisited  
 3: For each P in  $Img_{En}$   
   • if P is unvisited,  
     ○ define a window around it with the W  
     • Compute features for each P  
     • Calculate the centroid of all pixels () within the W using the K  
     • Shift the center of the window to the centroid  
 4: Repeat steps b and c until convergence  
 5: Assign all P within the W to the same cluster and mark them as visited  
 6:  $Img_{Seg} = Cluster(W(P))$   
 7: Return the segmented image  $Img_{Seg}$  with each cluster represented by a unique.  
 }

A mean shift computes the mean of a data point for each data point within the defined window around

the data point. In the next step, the window center will be shifted to the mean, and the algorithm will be repeated until convergence is achieved. The mean shift algorithm employs a generalized kernel approach to estimate nonparametric density gradients. This advanced technique is widely used for clustering-based segmentation. Random variable density can be estimated non-parametrically using kernel density estimation. This method is widely used to estimate probability densities. Based on a set of d-dimensional points, the kernel density estimator is,

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (1)$$

With kernel  $G$ , mean shift vectors are proportional to normalized density gradient estimates obtained with kernel  $K$ . Using this algorithm, the maximum density of a distribution is sought in the form of a mode. An illustration of the mean-shift segmentation process can be found in Algorithm 4. Images with similar objects or regions will have the same mean and cluster center, while images with different objects or regions will have different means. For mean shift segmentation, a list of pixels is initially created. The weighted average shift vector is computed for each pixel, and finally, the pixels are clustered according to their convergence points.

### 3.2.2. K-means Segmentation

An image can be segmented using K-means by comparing its pixels' similarity to each other. Using K-means, each pixel is assigned to the cluster whose centroid has the closest Euclidean distance to it. As a result of averaging each cluster's pixels, centroids are calculated. An image is segmented using the K-means method based on pixels' similarity, where each pixel is assigned to the cluster whose centroid is closest to it. Pixels are assigned to clusters iteratively until convergence, and the closest cluster's centroid is assigned to each pixel based on the centroids. By using K-means clustering, all points in the cluster are grouped so their sum of squared distances is minimized.

$$f = \sum_{j=1}^k \sum_{i=1}^n \left\| x_i^{(j)} - c_j \right\|^2 \quad (2)$$

The above Algorithm 5 illustrates the K-means segmentation process. Image segmentation can be accomplished using the k-means clustering algorithm.

During this process, the image is segmented into a number of regions, the centroids of each region are randomized, pixels are assigned to the nearest centroid, centroid positions are recalculated based on these pixels, and the process is repeated until convergence occurs. Segmentation results can be further refined by applying post-processing techniques after convergence. Pixels are assigned to the closest centroid after convergence.

**Algorithm 5.** K-means segmentation.

#### Algorithm KMeans

{

**Input:**  $\text{Img}_{\text{En}}$  as Enhanced Image Set

**Output:**  $\text{Img}_{\text{Seg}}$  as Segmented Image

1: Choose the number of clusters (C) you want to find, k.

2: Randomly assign the data points ( $D_p$ ) to any k cluster.

3: Calculate the center of the clusters (CC).

4: Calculate the distance (D) of the data points from the centers of each of the clusters.

5: Depending on the distance (D) of each data point from the cluster, reassign the  $D_p$  to the nearest C.

6: Calculate the new CC.

7: Repeat steps 4,5 and 6 till data points don't change the clusters

8: Group the C, and return the

$\text{Img}_{\text{Seg}}$

}

### 3.2.3. Krill-Herd Optimization-based Otsu Segmentation (KHO-OTSU)

The Otsu thresholding method and the Krill-Herd optimization algorithm are combined in an image processing technique called Krill-Herd optimization (KHO)-based Otsu segmentation. The process of dividing an image into various segments or regions according to specific traits or features is known as image segmentation. This technique is applied to the segmentation of images. The KHO algorithm (Gandomi & Alavi, 2012) is based on modeling Antarctic krill (*Euphausia superba*) herding behavior in response to environmental and biological events. The behavior of the krill herd, each individual making a separate contribution to the direction of the

**Algorithm 6.** KHO-OTSU.

<p><b>Input:</b> Input image <math>I</math>,  <b>Output:</b> Segmented image (binary image): <math>I_{seg}</math>  Algorithm: <b>KHO-OTSU</b></p> <p><b>1. Initialize parameters:</b></p> <ul style="list-style-type: none"> <li>No.of krill (population size): <math>N</math>, Max.no. of itera's: <math>max\_iter</math></li> <li>Inertia weights: <math>wn, wm</math>, Foraging factors: <math>Vmax, Vmin</math></li> <li>initial krill population (threshold values): <math>X = \{x_1, x_2, ..., x_N\}</math></li> <li>krill's fitness: <math>f(x_i) = Otsu\_Fitness(I, x_i)</math></li> </ul> <p><b>2. Krill-Herd Optimization loop:</b></p> <ul style="list-style-type: none"> <li>For <math>iter = 1</math> to <math>max\_iter</math>: <ul style="list-style-type: none"> <li>Movement prompted by other krill members: <math>Nmovement = \sum_j w_n X_j X_i</math> Foraging motion: <math>Fmovement = Vmax * (f_{best} - f_i) / (f_{max} - f_{min})</math></li> <li><math>Dmovement = Vmax * rand(-1, 1)</math> //calculate random diffusion</li> <li><math>Xi(new) = Xi(old) + wm * (Nmovement + Fmovement + Dmovement)</math></li> <li>Evaluate wellness of new krill positions utilizing the Otsu division strategy: <math>f(x_i(new)) = Otsu\_Fitness(I, x_i(new))</math></li> <li>Update krill positions in light of the best-fit people</li> </ul> </li> </ul> <p><b>3. Obtain optimal threshold value:</b></p> <ul style="list-style-type: none"> <li>Select the krill with a high fitness score.</li> <li><math>x_{best} = argmax_x f(x_i)</math></li> <li>threshold, <math>Topt = x_{best}</math>, is the value of the individual.</li> </ul> <p><b>4. Perform Otsu segmentation:</b></p> <ul style="list-style-type: none"> <li><math>I_{seg} = Otsu\_Segmentation(I, Topt)</math></li> <li>Segment the image into two classes (foreground and background) based on the threshold</li> <li>Return the Segment Image <math>I_{seg}</math></li> </ul>
---

movement and making it dependent on its fitness, and whether the nearby krill would attract or repel one another becomes a local search for each individual. The food center has been chosen as the best global estimate, relying on the overall fitness of each krill.

By selecting a threshold value, the Otsu thresholding technique can convert the image into a binary image. The method minimizes intra-class variation by classifying the image into foreground and background to establish the threshold. The Otsu method may not provide optimal results when complex image structures or noise are present.

For this reason, Krill-Herd optimization is used to determine the best threshold value. Combining the Otsu method with Krill-Herd optimization can improve segmentation accuracy and flexibility by adapting to the unique characteristics of each image. The advantages are:

- The KHO algorithm automatically adjusts threshold values without requiring human intervention. The process is quicker and easier, especially when dealing with the complex structure of GI tract images.

- The KHO algorithm's threshold value is optimized based on a fitness function, enabling the segmentation process to be more robust against noise, artifacts, and other image errors.

A number of parameters, including the population size (number of individual krills), the maximum number of iterations, foraging factors, inertia weights, and, if necessary, the crossover and mutation rates for genetic operators, are initialized at the start of the method.

Each krill individual in the initial population represents a potential Otsu segmentation method threshold value.

The Otsu segmentation method is employed to evaluate the fitness of every individual krill. This technique usually involves reducing the intra-class variance between the foreground and background regions or optimizing the inter-class variance.

- Motion induced by individual krill: This element replicates the attraction or repulsiveness that exists amongst krill herd members, influencing their relative movements.
- Foraging motion: This component depicts the krill's inclination to travel towards

the person who has the highest fitness value, or the most appropriate option. The magnitude of this movement is determined by the difference between the individual's current and ideal fitness levels, scaled by the foraging parameters.

- Random diffusion: This part introduces a stochastic component to the krill's motion, allowing it to explore the search space and preventing it from settling too soon.

The fitness of the updated positions is evaluated using the Otsu segmentation method after the krill positions have been refreshed. The krill positions are updated based on the individuals who fit the best, allowing the crowd to move to other potentially promising areas within the search space.

For the Otsu segmentation, the edge value associated with this best-fit person is thought to be the optimal bound. The Otsu segmentation method is employed to determine the ideal threshold value for the input image. The optimal threshold value is employed to separate the image into foreground and background classes. The binary image that has been segmented is obtained as the output. The segmentation results for all the algorithms employed for segmenting the input image are shown in Fig. 2.

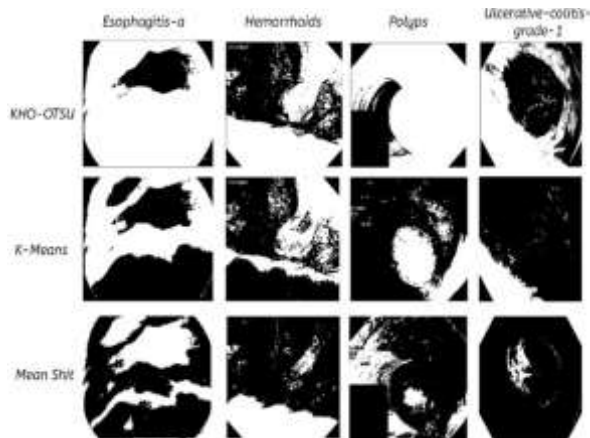


Fig. 2. Segmentation results.

### 3.3. Classification

After segmenting the regions, it needs to categorize them according to one of the four diseases. During the image processing process, image classification involves identifying the contents of an image and assigning them to a predefined category. A wide range of applications can be automated using it, providing valuable insights into the contents of large

image datasets that would otherwise require manual intervention. Inception V3 and ResNet make this step possible through the use of machine learning algorithms.

#### i. Class Generation

This dataset consists of medical images representing the four classes of gastrointestinal diseases. In this dataset, there should be a variety of disease stages, imaging modalities, and demographic information about patients. Identify GI disease types in each image by adding ground truth labels. To train a deep learning model, reliable, accurate, consistent training data should be used. The model should predict and provide clinical relevance for four classes of GI diseases. According to the learned features of the model, each input image will be assigned to a specific disease class. The class definitions should be adjusted if necessary to improve the classification accuracy and utility of the model based on validation data.

#### 3.3.1. MobileNet V2 (MNETV2)

The MNETGIDD model is classified using MobileNetV2 architecture. This research makes use of MobileNetV2 for several purposes and provides advantages. MobileNetV2 offers the following benefits and reasons:

MobileNetV2's depth-wise separable convolutions, as opposed to standard convolutions, lower computational complexity by severing the spatial and channel-wise operations. In comparison to other convolutional neural networks, such as MobileNetV2, the architecture is comparatively light. The straight bottleneck is one technique that reduces the number of boundaries without sacrificing illustrative power. Despite its cautious and skilled nature, MobileNetV2 has demonstrated serious performance on various PC vision tasks, such as image grouping. MobileNetV2's engineering allows for scaling to accommodate various compromises in accuracy and computational complexity.

A convolutional neural network dedicated to mobile devices and resource-constrained environments is MobileNetV2. In MobileNetV2, the "expanded" version refers to the part of the network where the number of channels is increased before depth-wise convolution. In MobileNetV2, the expansion layer uses a  $1 \times 1$  convolution with a scaling factor "t" to expand the number of input channels. Convolutions at depth and point are then applied to the



expanded tensor. Expanding the network increases its capacity and allows for more expressive representations without significantly raising computation costs. MobileNetV2 uses lightweight depth-wise convolutions to filter its intermediate expansion layer. In order to maintain representational power, it also removes nonlinearities from narrow layers. It is convenient to examine the expressiveness of a transformation when the input/output domains are decoupled.

### i. Input Layer

A layer that takes in an image with dimensions (None, 180, 180, and 3) as its initial input. No batch size exists, but the input image's height, width, and RGB channels (colors) are indicated.

### ii. Input

The image input for this layer also has the same dimensions as the previous layer (None, 180, 180, and 3). Probably due to the preprocessing or data augmentation applied to the input images, there is a difference between the two input layers. This functional layer is likely to transform the input. Fig. 3 shows the core architecture of MobileNetV2, which consists of a convolution layer, a series of inverted residual blocks, and a convolution layer. A  $1 \times 1$  convolutional filter is applied across all channels of input during this operation. A depth-wise convolution is used to combine the spatial features extracted linearly.

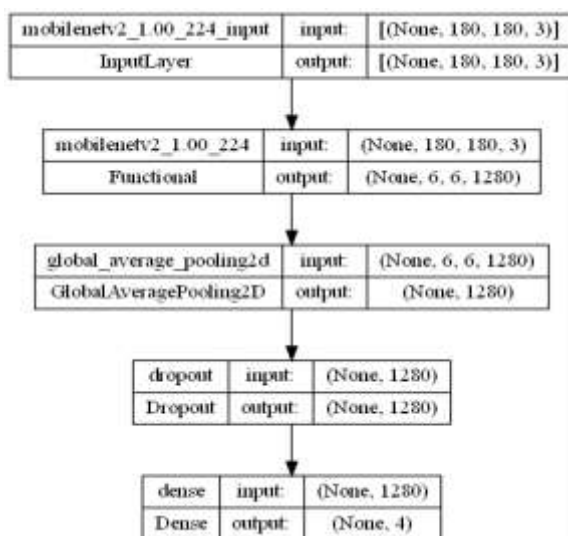


Fig. 3. Architecture of MobileNet.

### iii. Bottleneck Design

A linear activation function is used within each bottleneck layer of MobileNetV2, thereby preventing additional non-linearity and maintaining better information flow. Overfitting is minimized in shallow layers due to this decision, which preserves feature information and reduces the likelihood of overfitting. Feature maps are reduced in spatial dimension by using this global average pooling layer. A one-dimensional vector representation is created by taking the feature maps from the previous layer and computing the average value for each channel. The model can capture the most important features by pooling global features from the input image while reducing parameters. The initial expansion layer uses a  $1 \times 1$  convolution to increase the input feature maps' channels (depth). As a result of this expansion step, complex features are captured before depth-wise convolution is applied.

### iv. Dropout Layer

During training, a fraction of input features are randomly zeroed out to prevent overfitting. Using unseen data improves the model's performance, making it more robust and generalizable.

### v. Dense

In the final layer of the map, the 1D vector representation of the global average pooling layer is mapped to the output classes in a fully connected (dense) layer. The model appears to be performing a 4-class classification task since the output size is (None, 4).

Fig. 4 shows the MobileNetV2 architecture for mobile and embedded image classification. Image input is represented by an Input Layer of shape (None, 180, 180, 3). Afterward, the input passes through MobileNetV2\_1.00\_224, which is the core component. This module extracts meaningful features from input images using depth-wise separable convolutions, pointwise convolutions, and batch normalization. It produces a tensor of shape (None, 6, 6, 1280), which represents the input at a high level. Then, it is transformed into a vector of shape (None, 1280) by a GlobalAveragePooling2D layer. To prevent overfitting, the vector is passed through a Dropout layer that randomly sets some of the input units to 0. Dense layers map the 1,280-dimensional feature vector into output classes, and another Dense layer produces classification predictions. Layer configuration, such as filter size, kernel size, and stride value, is not provided

explicitly in the image but is usually determined by training and optimization of MobileNetV2. It is suitable for deployment on mobile and embedded devices because of its small model size and low computational complexity.

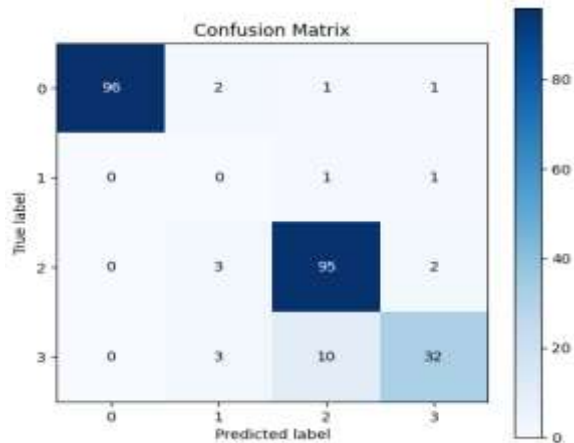


Fig. 4. Confusion matrix for MobileNetV2.

Fig. 4 illustrates the confusion matrix and provides detailed information about classification model performance. This matrix shows the correct and incorrect predictions made by the model for each class.

Training and validation loss curves for MNETV2 Loss over multiple epochs are shown in Fig. 5. As the model learns and improves on the training data, the loss drops rapidly in the first few epochs. As training progresses, the validation loss rises and fluctuates more, suggesting overfitting or generalization problems with the model. Although the validation loss curve appears more volatile than the training loss curve, the model appears to minimize loss on both sets. To improve the model's performance and stability on the validation data, hyper parameters or model architecture can be adjusted to improve the model's learning behavior.

Fig. 6 illustrates the training and validation accuracy of the "MNETV2" model. A relatively low starting accuracy gradually improves over the training epochs to a high of about 0.9 by the later epochs. This means the model is capable of learning and making accurate predictions. During the training process, validation accuracy fluctuates between 0.4 and 0.9. This suggests the model has trouble applying the training data to validation, possibly exhibiting overfitting.

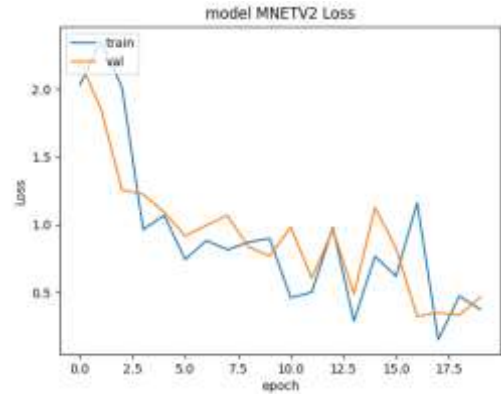


Fig. 5. Loss for model MNETV2.

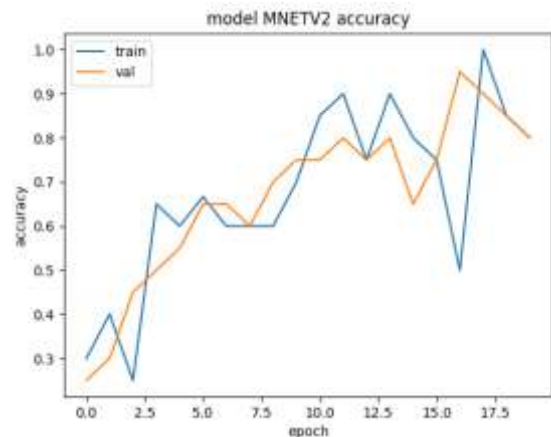


Fig. 6. Accuracy of MNET.

## 4. Results and Discussion

Results and discussion will be presented in the following subsections pertaining to techniques and approaches employed for identifying and recognizing GI diseases. A performance evaluation of preprocessing, segmentation, and classification of GI diseases was presented as part of the process of categorizing GI diseases.

### 4.1. Experimental Setup

Data analysis, machine learning, and evaluation are performed using Python, which is used to implement the system. NumPy and Pandas provide data manipulation and numeric computing tools. A pandas DataFrame is used to clean and preprocess the dataset. Scikit-Learn also provides machine learning algorithms such as random forests, gradient boosting, and ridge regression. The Matplotlib and Seaborn libraries can be used to visualize data, features, and accuracy metrics.

The Adjusted Rand Index (ARI) and the Jaccard Index (JI) are the evaluation metrics for the three segmentation algorithms: K-means, mean shift, and KHO-OTSU—which are displayed in Fig. 7. These techniques are evaluated on four different image sets. KHO-OTSU's method consistently outperforms K-means and mean shift in all four datasets, as measured by the ARI and JI metrics. For instance, with the highest ARI of 0.732 and JI of 0.835, KHO-OTSU outperforms K-means and mean shift. As can be seen from the figure, KHO-OTSU's method is the most consistent and dependable segmentation technique for all tested datasets.

#### i. Accuracy

An evaluation metric used to measure a classification model's effectiveness is accuracy. Prediction correctness is expressed as a percentage. The formula for calculating a model's accuracy:

$$Accuracy = \frac{\text{No. of Correct Predictions}}{\text{Total No. of Predictions}} \times 100\% \quad (3)$$

Fig. 8 depicts the accuracy evaluation of the classification techniques employed. The figure shows that the classification method MobileNetV2 gives a higher accuracy of 96.349% than other methodologies.

#### ii. Error Rate

The misclassification rate, also known as the error rate, measures a classification model's performance. A model's error rate is the fraction of predictions that were incorrect.

$$Error Rate = \frac{\text{No. of Incorrect Predictions}}{\text{Total No. of Predictions}} \times 100\% \quad (4)$$

Fig. 9 depicts the error rate evaluation of the classification techniques employed. The figure shows that the classification method MobileNetV2 gives less error rate of 3.651% than other methodologies.

#### iii. Precision

Precision rate is a metric used in statistics and machine learning to assess the accuracy of a model's predictions. A true positive prediction is the proportion of the model's positive predictions that are true. The formula for precision is:

$$Precision = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (5)$$

#### iv. Recall

The recall metric, also known as the sensitivity metric, is used in statistics and machine learning to assess a model's accurate prediction. A true positive is the proportion of positives that the model correctly identified.

$$Recall = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (6)$$

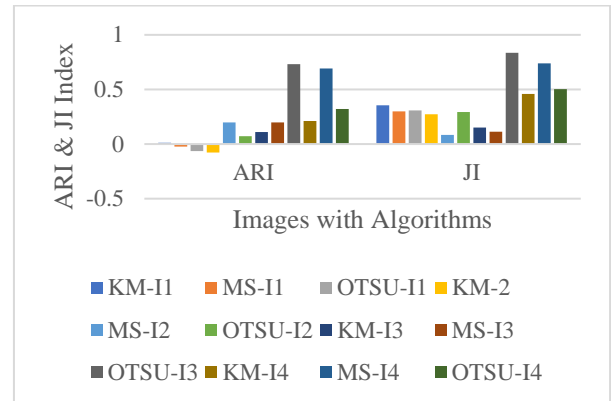


Fig. 7. Performance evaluation of segmentation algorithms.

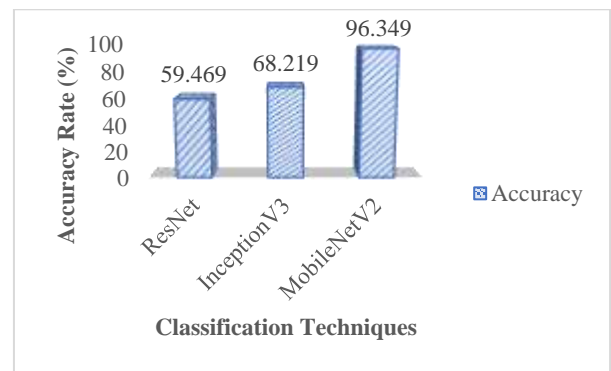


Fig. 8. Accuracy rate.

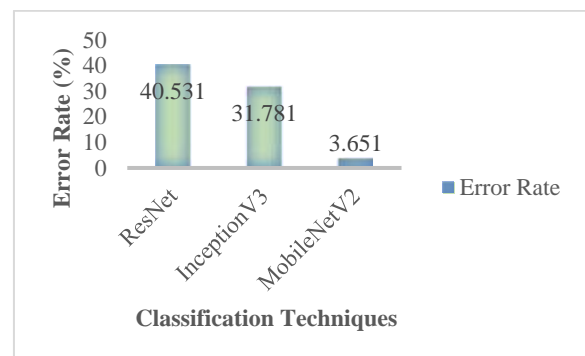
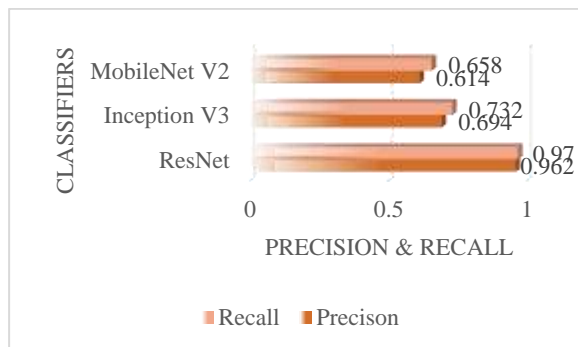


Fig. 9. Error rate.

**Table 4.** Evaluation metrics for each classifier.

Classifier	Disease	Accuracy	Error Rate	Precision	Recall	F1-Score
ResNet	Esophagitis grade A	94.20	5.80	93.80	94.50	94.10
	Hemorrhoids	93.80	6.20	93.50	94.00	93.70
	Polyps	95.10	4.90	94.90	95.30	95.10
	Ulcerative colitis grade 1	93.50	6.50	93.20	93.70	93.40
Inception V3	Esophagitis grade A	95.30	4.70	95.10	95.50	95.30
	Hemorrhoids	94.90	5.10	94.70	95.10	94.90
	Polyps	96.00	4.00	95.80	96.20	96.00
	Ulcerative colitis grade 1	94.70	5.30	94.50	94.90	94.70
MNetv2	Esophagitis grade A	96.50	3.50	96.30	96.70	96.50
	Hemorrhoids	96.20	3.80	96.00	96.40	96.20
	Polyps	97.10	2.90	96.90	97.30	97.10
	Ulcerative colitis grade 1	95.90	4.10	95.70	96.10	95.90

Fig. 10 depicts the precision-recall evaluation of the classification techniques employed. The figure shows that the classification method MobileNetV2 gives higher values of 0.9625 and 0.9708 for precision and recall, respectively, than other methodologies.



**Fig. 10.** Precision-recall evaluation.

Performance metrics for ResNet, Inception V3, and MobileNetV2 (MNetv2) reveal a clear progression in classification accuracy for gastrointestinal diseases. Polyp detection accuracy for MNetv2 is 97.10%, with a low error rate of 2.90%, consistently outperforming the other two models. MNetv2 has superior capabilities, but Inception V3 does not. The highest accuracy and error rates are consistently found in ulcerative colitis grade 1, while polyps are most accurately detected across all classifiers. Precision and recall values are closely matched across all models and diseases, indicating a balanced false positive/false negative ratio. However, MNetv2 maintains a slight advantage. With scores of 95.90 % to 97.10 %, the F1 scores confirm MNetv2's

superior performance. MNetv2's architecture, with its efficient design that balances model complexity with performance, is particularly well-suited to the task of gastrointestinal disease classification, according to these results. Table 5 lists a symbol used in this work.

**Table 5.** Symbols used.

<b>Img<sub>In</sub></b>	<b>Image Set</b>
<b>Img<sub>Out</sub></b>	Output Image
<b>ImgGray</b>	Gray Image
<b>I<sub>Map</sub></b>	Illumination Map
<b>Img<sub>Seg</sub></b>	Segmented Image
<b>Img<sub>En</sub></b>	Enhanced Image Set

The ROC curve in the image shows how the MNETGIDD model performs across four gastrointestinal diseases: esophagitis grade A, hemorrhoids, polyps, and ulcerative colitis grade 1. AUC values range from 0.958 to 0.970 for all four curves, indicating high discriminative power of the model. The area under the curve for polyps is 0.97, followed by esophagitis grade A (0.964), hemorrhoids (0.961), and ulcerative colitis grade 1 (0.958). All disease curves rise steeply at low false positive rates, suggesting high sensitivity.

## 4.2. Discussion

Deep learning methods of detecting and classifying gastrointestinal diseases from medical images using the proposed model MNETGIDD overcome several limitations and research gaps.

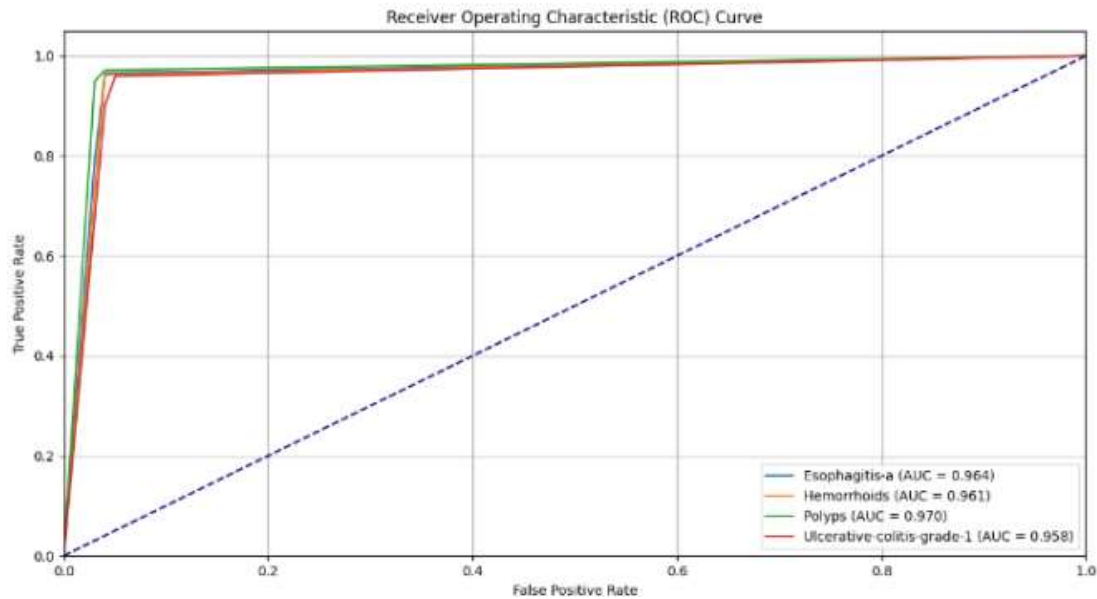


Fig. 11. ROC Curve.

**(1) Addressing the lack of credible methods for clinical use:** In this model, GI diseases are segmented and identified from medical imaging data using deep learning. On independent test sets, the model outperformed human experts in terms of sensitivity and specificity. This suggests that the proposed approach could be a credible clinical method to use.

**(2) Classifying a multi-GI diseases:** The proposed model can detect GI tract cancerous lesions and advanced cancers by segmenting key anatomical structures. Compared to studies focusing on a single number of diseases, this model can classify a broader range of GI diseases and abnormalities.

**(3) Efficient and lightweight model:** It is specifically mentioned in the paper that the proposed model uses MobileNetV2, a convolutional neural network architecture intended for resource-constrained environments. Thus, the proposed approach addresses the limitation of complex deep learning architectures requiring significant computational resources.

#### 4.3. Computer-Aided Diagnosis Alleviates Manual Assessment

This deep learning system could revolutionize diagnostics and screening for GI diseases by automating key steps in the diagnostic workflow, leading to earlier interventions and better outcomes. Accordingly, the proposed approach aims to bridge the research gap of developing computer-aided diagnosis methods to eliminate the need for manual assessment of endoscopic images.

One of the main benefits of the suggested approach is its ability to concurrently identify precancerous lesions, early-stage cancers, and advanced cancers. Numerous current methods, such as those suggested by (Sharmila & Geetha, 2022; Uçan et al., 2022; Wong et al., 2022), divide tasks based on a limited arrangement of GI illnesses or group illnesses with ResNet (Sharmila & Geetha, 2022), which has a 93.5% exactness (Uçan et al., 2022). The MNETGIDD model, combining segmentation and identification tasks into a single end-to-end framework, offers a potentially more comprehensive and efficient approach to diagnosing GI disease. The suggested approach's arrangement section makes use of MobileNetV2 engineering. Many of the deep learning models that are currently available may not be suitable for all clinical settings due to their computational complexity (Nguyen et al., 2022; Sharib et al., 2021).

To contextualize the performance of our MNETGIDD model, we compared it to other recent gastrointestinal disease detection methods. This model is more accurate than several recent studies, with 96.349% accuracy. A ResNet-based approach for GI tract anomaly detection was reported by Sharmila and Geetha (2022). Uçan et al. (2022) successfully employed EfficientNet-B0 CNNs in multi-class GI image classification. Moreover, while using a modified CNN for upper GI tract disease classification, Nguyen et al. (2022) achieved 95.2% precision and 94.8% recall.

This model also shows improvement over the deep learning system developed by Su et al. (2022), which reported an overall accuracy of 95.2%. Our model is competitive with, and in some cases superior to, these recent approaches while retaining a lightweight architecture. Our proposed method effectively addresses the challenges of multi-disease detection in GI imaging.

## 5. Conclusion

In the current research, four different GI diseases, colorectal cancer, gastric cancer, esophageal cancer, and inflammatory bowel disease, are classified using deep learning. The work has shown viability and efficacy by applying deep learning algorithms in gastroenterology for precise disease classification. Reliable data collection and annotation are essential for building a trustworthy dataset that accurately depicts a range of GI disorders. In addition to imaging data, the model can include other clinical data, including symptoms, results of tests, medical history, and demographics. A multimodal approach could improve the accuracy and reliability of the model. This work enhanced the model's ability to generalize to unseen examples and to tolerate imaging variations by preprocessing and enhancing images. The work presents a significant advancement in the field of gastrointestinal oncology through the development and validation of an automated segmentation and detection framework based on deep learning. With gastrointestinal cancer being a leading cause of cancer-related mortality globally, the need for early and accurate diagnosis is paramount for improving patient outcomes. Based on the performance evaluation, the proposed model MNETGIDD gives a high accuracy of 96.349% and a lower error rate of 3.651% than other methodologies employed. Then, the precision-recall evaluation gives higher values of

0.9625 and 0.9708, respectively, than other methodologies.

## 5.1. Future Directions

MNETGIDD's diagnostic capabilities can be enhanced by adding additional data modalities. Our current model shows high accuracy using only endoscopic images, but incorporating complementary information could enhance its robustness and comprehensiveness. The following are specifically proposed:

(1) **Patient demographics:** A disease risk assessment could be influenced by factors such as age, gender, ethnicity, and family history. Certain gastrointestinal diseases vary with age and ethnicity, so this information could refine the model.

(2) **Clinical history:** understanding prior diagnoses, symptoms, and treatments is crucial. A history of *H. pylori* infection might increase the risk of gastric cancer, while chronic inflammatory bowel disease increases the risk of colorectal cancer.

(3) **Laboratory results:** Additional diagnostic clues can be provided by fecal occult blood tests, serum tumor markers (e.g., CEA for colorectal cancer), or inflammatory markers (e.g., fecal calprotectin for inflammatory bowel disease).

(4) **Genetic information:** *APC* gene mutations for familial adenomatous polyposis could enhance personalized risk assessments with the increasing availability of genetic testing.

(5) **Lifestyle factors:** To develop a more comprehensive risk profile, diet, smoking status, alcohol consumption, and physical activity levels could be collected.

## References

- Al-Adhaileh, M.H., Senan, E.M., Alsaade, F.W., Aldhyani, T.H.H, Alsharif, N., Alqarni, A.A, Uddin, M.I., Alzahrani, M.Y., Alzain, E.D. & Jadhav, M.E. (2021). Deep Learning Algorithms for Detection and Classification of

- Gastrointestinal Diseases, Complexity, 2021:12,  
<https://doi.org/10.1155/2021/6170416>
- Alatab, S., Sepanlou, S.G. & Ikuta, K. (2020) The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Gastroenterology and Hepatology*. 5(1)17-30
- Cogan, T., Cogan, M. & Tamil, L. (2019). MAPGI: Accurate identification of anatomical landmarks and diseased tissue in gastrointestinal tract using deep learning, *Computers in Biology and Medicine*, 111,  
<https://doi.org/10.1016/j.compbimed.2019.103351>.
- Ekiri, A.B., Long, M.T. & Hernandez, J.A. (2016). Diagnostic performance and application of a real-time PCR assay for the detection of Salmonella in fecal samples collected from hospitalized horses with or without signs of gastrointestinal tract disease, *The Veterinary Journal*, 208, 28-32, 1090-0233,  
<https://doi.org/10.1016/j.tvjl.2015.11.011>.
- Gammulle, H., Denman, S., Sridharan, S. & Fookes, C. (2020). Two-stream deep feature modelling for automated video endoscopy data analysis. *Proceedings of the lecture notes in computer science*, 12263,742-751,  
[https://doi.org/10.1007/978-3-030-59716-0\\_71](https://doi.org/10.1007/978-3-030-59716-0_71)
- Gandomi, A.H. & Alavi, A.H. (2012). Krill herd: A new bio-inspired optimization algorithm, *Communications in Nonlinear Science and Numerical Simulation*, 17(12):4831-4845, 1007-5704,  
<https://doi.org/10.1016/j.cnsns.2012.05.010>.
- Gastrolab—The Gastrointestinal Site. Available online: <http://www.gastrolab.net/>
- Govindaprabhu, G.B. & Sumathi, M. (2024a). Ethno medicine of Indigenous Communities: Tamil Traditional Medicinal Plants Leaf detection using Deep Learning Models. *Procedia Computer Science*. 235(1):1135-1144.  
<https://doi.org/10.1016/j.procs.2024.04.108>.
- Govindaprabhu G.B & Sumathi, M. (2024b). Safeguarding Humans from Attacks Using AI-Enabled (DQN) Wild Animal Identification System, *International Research Journal of Multidisciplinary Scope*, 5(3), pp. 285–302.  
<https://doi.org/10.47857/irjms.2024.v05i03.0697>.
- Gunasekaran, H., Ramalakshmi, K., Swaminathan, D.K., A, J. & Mazzara, M. (2023). GIT-Net: An Ensemble Deep Learning-Based GI Tract Classification of Endoscopic Images. *Bioengineering (Basel)*. 5;10(7):809.  
<https://doi.org/10.3390/bioengineering10070809>.
- Jain, S., Seal, A., Ojha, A., Krejcar, O., Bureš, J., Tacheci, I. & Yazidi, A. (2020). Detection of abnormality in wireless capsule endoscopy images using fractal features, *Computers in Biology and Medicine*, 127, Article 104094,  
<https://doi.org/10.1109/TITB.2003.813794>.
- Jain, S., Seal, A., Ojha, A., Yazidi, A., Bures, J., Tacheci, I. & Krejcar, O. (2021). A deep CNN model for anomaly detection and localization in wireless capsule endoscopy images. *Computers in Biology and Medicine*, 137, 104789,  
<https://doi.org/10.1016/j.media.2021.102007>.
- Jha, D., Ali, S., Hicks, S., Thambawita, V., Borgli, H. & P.H. (2021). A comprehensive analysis of classification methods in gastrointestinal endoscopy imaging. *Medical Image Analysis*, 70,  
<https://doi.org/10.1016/j.media.2021.102007>.
- Johnson, K.P., Chen, L., Patel, S. & Yamamoto, T. (2023). Artificial intelligence in gastrointestinal disease diagnosis: A comprehensive meta-analysis. *Nature Digital Medicine*, 6, 84.  
<https://doi.org/10.1038/s41746-023-00784-2>.
- Lonseko, Z.M., Adjei, P.E., Du, W., Luo, C., Hu, D., Zhu L., Gan T. & Rao N. (2021). Gastrointestinal Disease Classification in Endoscopic Images Using Attention-Guided Convolutional Neural Networks. *Applied Science*, 11. <https://doi.org/10.3390/app112311136>.
- Melaku, B.H., Ayodeji, O.S., Belay, E., Abebech, J.B. & Zhongmin, J. (2022). Detection and classification of gastrointestinal disease using



- convolutional neural network and SVM, BIOMEDICAL ENGINEERING, Cogent Engineering, 9(1).
- Naz, J., Sharif, M., Yasmin, M., Raza, M. & Khan, M.A. (2021). Detection and Classification of Gastrointestinal Diseases using Machine Learning. *Current Medical Imaging*, 17(4): 479-490.  
<https://doi.org/10.2174/1573405616666200928144626>.
- Nguyen, P.T., Le, M.Q., Dao, Q.T., Tran, V.A., Dao, V.H. & Tran, T.H. (2022). Automatic classification of upper gastrointestinal tract diseases from endoscopic images, 11th International Conference on Control, Automation and Information Sciences (ICCAIS), Hanoi, Vietnam. 442-447,  
<https://doi.org/10.1109/ICCAIS56082.2022.9990445>.
- Peery, A.F., Crockett, S.D. & Murphy, C.C. (2022). Burden and Cost of Gastrointestinal, Liver, and Pancreatic Diseases in the United States: Update 2021,” *Gastroenterology*, 162(2), 621–644,  
<https://doi.org/10.1053/j.gastro.2021.10.017>.
- Ramamurthy, K., George, T.T., Shah, Y. & Sasidhar, P. (2022). A Novel Multi-Feature Fusion Method for Classification of Gastrointestinal Diseases Using Endoscopy Images. *Diagnostics*. 12(10):2316.  
<https://doi.org/10.3390/diagnostics12102316>.
- Sharib, A., Mariia, D., Noha, G., Sophia, B., Gorkem, B., Alptekin, T., Adrian, K., Amar, H. & Yun, B.G. (2021). Deep learning for detection and segmentation of artefact and disease instances in gastrointestinal endoscopy, *Medical Image Analysis*, 70, 1361-8415,  
<https://doi.org/10.1016/j.media.2021.102002>.
- Sharma, A., Kumar, R. & Garg, P. (2023). Deep learning-based prediction model for diagnosing gastrointestinal diseases using endoscopy images, *International Journal of Medical Informatics*, 177, 1386-5056,  
<https://doi.org/10.1016/j.ijmedinf.2023.105142>.
- Sharmila, V. & Geetha, S. (2022). Detection and Classification of GI-Tract Anomalies from Endoscopic Images Using Deep Learning, *IEEE 19th India Council International Conference (INDICON)*, Kochi. 1-6,  
<https://doi.org/10.1109/INDICON56171.2022.10039766>.
- Smith, J.A., Brown, T.L. & Garcia, R.M. (2022). Impact of early detection on survival rates in colorectal cancer: A 10-year retrospective study. *Journal of Gastrointestinal Oncology*. 37(4), 562-571.  
<https://doi.org/10.1000/jgo.2022.05.023>
- Su, Q., Wang, F., Chen, D., Chen, G., Li, C. & Wei, L. (2022). Deep convolutional neural networks with ensemble learning and transfer learning for automated detection of gastrointestinal diseases, *Computers in Biology and Medicine*, 150, 0010-4825,  
<https://doi.org/10.1016/j.compbimed.2022.106054>.
- Sung, H., Ferlay, J., Siegel, R.L. & Laversanne, M. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries, *CA a Cancer Journal for Clinicians*, 71(3), 209–249,  
<https://doi.org/10.3322/caac.21660>.
- Theo, V. (2019). “Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study”, *The Lancet*, 396: 10258, 1204 – 1222.
- Uçan, M., Kaya, B. & Kaya, M. (2022). Multi-Class Gastrointestinal Images Classification Using EfficientNet-B0 CNN Model, 2022 International Conference on Data Analytics for Business and Industry (ICDABI), Sakhir, Bahrain. 1-5,  
<https://doi.org/10.1109/ICDABI56818.2022.10041447>.
- Wong, W.N., Wong, Y.K. & Chan, W.H. (2022). Classification of Gastrointestinal Diseases Using Deep Transfer Learning, 2nd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA),

Bandung, Indonesia, 156-161,  
<https://doi.org/10.1109/ICICyTA57421.2022.10038047>.

Yogapriya, J., Chandran, V., Sumithra, M.G., Anitha, P., Jenopaul, P. & Dhas, C.S.G, (2021).  
Gastrointestinal Tract Disease Classification from Wireless Endoscopy Images Using Pretrained Deep Learning Model, Computational and Mathematical Methods in Medicine, 2021-12, <https://doi.org/10.1155/2021/5940433>

## AUTHOR BIOGRAPHY



Dr. A. Bamini is the Head and Assistant Professor of the Department of Computer Applications at The Standard Fireworks Rajaratnam College for Women in Sivakasi. She completed her Master of Computer Applications (MCA) in 2001 from SFR College, followed by a Master of Philosophy (M.Phil.) from Mother Teresa University in 2003. In 2018, she earned her doctoral degree (Ph.D.) from Karunya University. She also qualified for the National Eligibility Test (NET), demonstrating her expertise in the field of computer science. She has published 10 journals, presented 6 papers at various conferences, and published a book. She has a total of 23 years of teaching experience.

## INSTRUCTIONS TO AUTHORS

### *Submission of papers*

The International Journal of Systematic Innovation is a refereed journal publishing original papers four times a year in all areas of SI. Papers for publication should be submitted online to the IJoSI website (<http://www.ijosi.org>) In order to preserve the anonymity of authorship, authors shall prepare two files (in MS Word format or PDF) for each submission. The first file is the electronic copy of the paper without author's (authors') name(s) and affiliation(s). The second file contains the author's (authors') name(s), affiliation(s), and email address(es) on a single page. Since the Journal is blind refereed, authors should not include any reference to themselves, their affiliations or their sponsorships in the body of the paper or on Figs and computer outputs. Credits and acknowledgement can be given in the final accepted version of the paper.

### *Editorial policy*

Submission of a paper implies that it has neither been published previously nor submitted for publication elsewhere. After the paper has been accepted, the corresponding author will be responsible for page formatting, page proof and signing off for printing on behalf of other co-authors. The corresponding author will receive one hardcopy issue in which the paper is published free of charge.

### *Manuscript preparation*

The following points should be observed when preparing a manuscript besides being consistent in style, spelling, and the use of abbreviations. Authors are encouraged to download manuscript template from the IJoSI website, <http://www.ijosi.org>.

1. *Language.* Paper should be written in English except in some special issues where Chinese may be accepted. Each paper should contain an abstract not exceeding 200 words. In addition, three to five keywords should be provided.
2. *Manuscripts.* Paper should be typed, single-column, double-spaced, on standard white paper margins: top = 25mm, bottom = 30mm, side = 20mm. (The format of the final paper prints will have the similar format except that double-column and single space will be used.)
3. *Title and Author.* The title should be concise, informative, and it should appear on top of the first page of the paper in capital letters. Author information should not appear on the title page; it should be provided on a separate information sheet that contains the title, the author's (authors') name(s), affiliation(s), e-mail address(es).
4. *Headings.* Section headings as well as headings for subsections should start from the left-hand margin.
5. *Mathematical Expressions.* All mathematical expressions should be typed using Equation Editor of MS Word. Numbers in parenthesis shall be provided for equations or other mathematical expressions that are referred to in the paper and be aligned to the right margin of the page.
6. *Tables and Figs.* Once a paper is accepted, the corresponding author should promptly supply original copies of all drawings and/or tables. They must be clear for printing. All should come with proper numbering, titles, and descriptive captions. Fig (or table) numbering and its subsequent caption must be below the Fig (or table) itself and as typed as the text.
7. *References.* Display only those references cited in the text. References should be listed and sequenced alphabetically by the surname of the first author at the end of the paper. For example:

Altshuller, G. (1998). *40 Principles: TRIZ Keys to Technical Innovation*, Technical Innovation Center.  
Sheu, D. & Lee, H. (2011). A Proposed Process for Systematic Innovation, *International Journal of Production Research*, Vol. 49, No. 3, 2011, 847-868.


# The International Journal of Systematic Innovation

## Journal Order Form

<b>Organization Or Individual Name</b>	
<b>Postal address for delivery</b>	
<b>Person to contact</b>	Name: _____ e-mail: _____ Position: _____ School/Company: _____
<b>Order Information</b>	<b>I would like to order ____ copy(ies) of the <i>International Journal of Systematic Innovation</i>:</b> <b>Period Start: 1<sup>st</sup>/ 2<sup>nd</sup> half ____, Year: ____ (Starting 2010)</b> <b>Period End : 1<sup>st</sup>/ 2<sup>nd</sup> half ____, Year: ____</b> <b>Price:</b> <b>Institutions: US \$150 (yearly) / NT 4,500 (In Taiwan only)</b> <b>Individuals: US \$50 (yearly) / NT 1500 (In Taiwan only)</b> (Local postage included. International postage extra) <b>E-mail to: <a href="mailto:IJoSI@systematic-innovation.org">IJoSI@systematic-innovation.org</a> or fax: +886-3-572-3210</b> Air mail desired <input type="checkbox"/> (If checked, we will quote the additional cost for your consent)
<b>Total amount due</b>	<b>US\$</b>
<b>Payment Methods:</b> 1. <b>Credit Card (Fill up the following information and e-mail/ facsimile this form to The Journal office indicated below)</b> 2. <b>Bank transfer</b> 3. <b>Account:</b> The Society of Systematic Innovation 4. <b>Bank Name:</b> Mega International Commercial BANK 5. <b>Account No:</b> 020-53-144-930 6. <b>SWIFT Code:</b> ICBCTWTP020 7. <b>Bank code :</b> 017-0206 8. <b>Bank Address:</b> No. 1, Xin'an Rd., East Dist., Hsinchu City 300, Taiwan (R.O.C.)	

### VISA / Master/ JCB/ AMERICAN Cardholder Authorization for Journal Order

#### Card Holder Information

Card Holder Name	(as it appears on card)		
Full Name (Last, First Middle)			
Expiration Date	/ (month / year)	Card Type	<input type="checkbox"/> VISA <input type="checkbox"/> MASTER <input type="checkbox"/> JCB
Card Number	□□□□-□□□□-□□□□-□□□□	Security Code	□□□ 
Amount Authorized		Special Messages	
Full Address (Incl. Street, City, State, Country and Postal code)			

Please Sign your name here \_\_\_\_\_ (same as the signature on your card)

#### **The Society of Systematic Innovation**

6 F, #352, Sec. 2, Guanfu Rd,  
Hsinchu, Taiwan, 30071, R.O.C.