International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
***Article Received:** 25 November 2024 **Revised:** 12 December 2024 **Accepted:** 30 January 2025*

PolypNet: A Lightweight CNN Framework for Early Detection of Colorectal Polyps Using Deep Learning

Nishat Tasnim
Department of Computer science and engineering
Bangladesh Army International University of Science and Technology
Cumilla, Bangladesh
nishattasnim565312@gmail.com

Ahmed Mamun
Faculty of Computer science and engineering, Department of Computer science and engineering
Bangladesh Army International University of Science and Technology
Cumilla, Bangladesh
mamun.cse@baiust.ac.bd

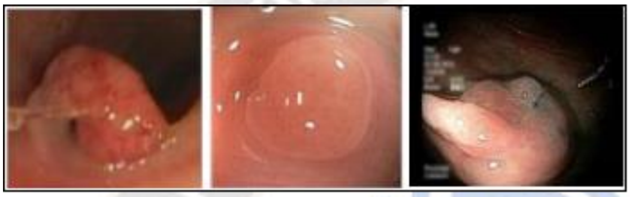
Md. Imrul Kayes
Department of Computer science and engineering
Chittagong University of Engineering and Technology
Cumilla, Bangladesh
md.imrulkayes6022@gmail.com

Abstract— Colorectal carcinoma is one of the most common reasons for carcinogenic death in the current world. Identifying the polyps that are present in the colon walls is one method to prevent this illness. However, a sparse number of research studies have been done to create a computer system that will detect the indisposition in the earlier stage. The enlargement of computer vision technology has accelerated the process by retrieving helpful information from the correlated data. Nonetheless, it is important to create an untrammelled system that will be able to sport colon polyps with better accuracy and training cost. In this research, we have delineated a Convolutional Neural Network (CNN) to emphasise Adenomatous, Hyperplastic and Serrated Lesions. The experiment of the network on the basis dataset has achieved an accuracy of 99.95% within a training time of only 18 minutes and 59 seconds. Stable learning efficiency was attained by the six-layer CNN with max-pooling and dropout regularisation.

Keywords- *Colorectal Cancer Detection, Polyp Classification, Convolutional Neural Network (CNN), Deep Learning, Data Preprocessing, Image Augmentation, State-of-the-Art , F1 Score and AUC.*

I. INTRODUCTION

Colorectal adenocarcinoma is the third leading cause of morbidity in both males and females worldwide. Statistics indicate that colorectal cancer could culminate in approximately 52,980 fatalities by 2021. According to the American Cancer Society's report, males are at a 4.3% lifetime risk of succumbing from colorectal cancer, while females are at a 4.0% risk [1]. When cells in the colon or rectum proliferate excessively and disseminate into adjacent organs, a specific type of cancer known as colon cancer is the result. Lifestyle choices, diabetes, obesity, genetic predispositions, and chronic inflammatory disorders are among the factors that increase the likelihood of cancer in the elderly. Among the available treatments include surgery, medicine and radiation therapy. Typically benign, colon polyps are growths that develop on the membrane of the colon. Colonoscopies and other routine screening technologies are indispensable for the early detection and safe removal of these lesions. In order to lower the number of colorectal cancer fatalities, colon polyps must be promptly detected and removed since they are a crucial indicator of the development of colon cancer.The most prevalent type of polyps is Adenomatous and they are frequently examined for malignant changes. Hyperplastic lesion sare frequently benign and are less likely to develop into cancer due to their low potential for malignancy. In the context of colorectal cancer, Serrated Lesions a collection of colorectal polyps with variable propensity for malignant transformation are significant due to their unique carcinogenic route from the conventional Adenoma carcinoma sequence.



Adenomatous Hyperplastic Serrated Lesions

Figure 1: Polyp Images

This route, which includes several subtypes such as Traditional Serrated Adenomas (TSA), Hyperplastic Polyps (HP) and Sessile Serrated Polyps/Lesions (SSP/SSL), is responsible for up to 15–30% of colorectal cancer cases. The research delineates a novel method for the identification of colon polyps, with a particular emphasis on those that are Adenomatous, Hyperplastic and contain Serrated Lesions. This study compared a multilayer Convolutional Neural Network (CNN) model to ten pre-trained Deep Learning State-of-the-Art (SOTA) models. The models included VGG-16, VGG-19, MobileNet, DenseNet-201, DenseNet-121, Inception-ResNet-V2, Inception-V3, Xception, EfficiencyNet B7 and ResNet-152-V2. The proposed CNN model prevailed over all of these architectures, obtaining the utmost accuracy of **99.95%** with a training duration of only **18 minutes and 59 seconds**. Deep learning, a subfield of artificial intelligence and machine learning, employs multi-layered neural network models to extract patterns and representations from extensive datasets. CNNs are particularly well-suited for image analysis due to their ability to extract spatial hierarchies from data. Xception, EfficientNet B7, ResNet-152-V2, MobileNet, DenseNet-201, DenseNet-121, Inception-ResNet-V2, Inception-V3, VGG-16 and VGG-19 are the SOTA models that are being compared. Since the CNN model offers a more economical and comprehensible architecture while preserving exceptional speed, it may be able to address some of the issues with SOTA models. Our research prioritises the user's experience, accuracy and computation efficacy in order to make a significant contribution to the development of medical imaging technology. Additionally, it underscores the importance of AI in improving the results of early detection and treatment for potentially fatal diseases such as colon cancer.

II. LITERATURE REVIEW

Md. Imrul Kayes, Rashida Feroz Prome, Maria Noor, Shovan Bhowmik, and Mamun Ahamed implemented a stacked convolutional neural network for colorectal cancer early detection. They classified "Adenomatous" and "Hyperplastic" polyps with a confidence level of up to 97.6%. The HOG feature detection approach was also implemented in the study, contributing to a substantial decrease in training time when juxtaposed with baseline models. In order to assess the models, cross-validation was implemented. Nevertheless, the research is limited by a variety of factors, such as the possible impacts of the dataset's imbalance, ambiguity in performance on other datasets and a lack of explanation for the model's interpretability, which is essential in the field of medicine. [1]

Study, Younas, Usman, and Yan categorized colon polyps determined through narrowband imaging endoscopy as Hyperplasic, Adenoma or Serrated. In order to analyze colon endoscopic images, they outlined a deep convolutional neural network (CNN) based weighted ensemble classification approach. Data is correctly classified by integrating ineffective learners' assets. The authors evaluated six distinct CNNs and implemented transfer learning to determine the optimal ensemble model architecture. The proposed approach outperformed the current models in terms of model dependability, F1-score, precision, recall, and accuracy. Meanwhile, the method's practical applicability may be restricted by its evaluation on two distinct datasets, the large training data and the high computational costs.[2]

Pallabi Sharma came up with LPNet, a convolutional neural network (CNN) that is both lightweight and efficient. It is specifically designed to aid in the classification of colon lesions from colonoscopy video frames. LPNet achieves an accuracy of 93.55% in polyp classification tasks by streamlining identification through the substitution of discrete wavelet transformation (DWT) pooling for max and average pooling. By incorporating the Gaussian noise convolution block, the model's ability to withstand fluctuations in colonoscopy frames is enhanced. Nevertheless, the model's efficacy may be limited by the quality of colonoscopy images and the absence of annotated datasets for colorectal cancer studies.[3]

Sushama Tanwar and other researchers formed a deep learning model to recognize and categorize colorectal lesions in colonoscopy images. Aiming to enhance the quality of images, the model implements dynamic histogram equalization and guided image filters. The rapid detection and classification of polyps is facilitated by the Single Shot MultiBox Detector (SSD). Despite its 92% accuracy rate, the model has not been assessed on large datasets and may not be appropriate for real-time applications due to its high computational complexity.[4]

Win Sheng Liew has devised a novel approach to the identification of colonic polyps that employs AI-based CAD tools. AdaBoost ensemble learning, principal component analysis and modified deep residual networks comprise the methodology. Its Matthews Correlation Coefficient of 98.13% serves as evidence of its exceptional precision, sensitivity, specificity, and accuracy. Nevertheless, it is costly, inefficient, and may obstruct discovery as a result of its limited ability to identify specific polyp types. Additional research and verification are required for clinical implementation.[5]

Debesh Jha and collaborators utilize deep learning to identify, localize, and segment colonoscopy lesions. The objective is to decrease the occurrence of undetected lesions during colonoscopy, which has the potential to result in colorectal cancer. On the Kvasir-SEG dataset, YOLOv4 with Darknet53 was the best accurate detection and localization model. For segmentation, ColonSegNet was suggested, with competitive performance metrics. Nevertheless, the investigator encounters a wide range of obstacles, such as data loss during photo resizing, challenges with sessile or flat polyps and challenges with retrospective design. [6]

Kangkana Bora employs automated feature analysis to classify colonic lesions as either neoplastic or non-neoplastic. Generic Fourier descriptors and contourlet transformations are filtered to extract shape information. The statistical significance of descriptor contributions is determined through the use of ANOVA. Final descriptors are optimized using fuzzy entropy-based feature ranking. They classify using Least Square Support Vector Machine and Multi-layer Perceptron. The research demonstrates that automated feature analysis can assist in the early detection of colorectal cancer, as evidenced by an average accuracy of 93.33% and an F-measure of 95.67%. Furthermore, the paper's purview is restricted by the restrictive comparison with deep learning models, the reliance on only two datasets, and the feature-based methodologies.[7]

III. DATASET OVERVIEW

Complete analysis is done using [7] to integrate public and private datasets. The Getafe University Hospital generously shared the dataset. Polyp videos are "Adenoma Lesions", "Serrated Lesions" and "Hyperplastic Lesions". Every category features White Light and Narrow Band Image Videos. Thirty "Serrated Lesions" (15 White Light and 15 Narrow Band Image), eighty "Adenoma Lesions" (40 White Light and 40 Narrow Band Image) and forty-two "Hyperplastic Lesions" (21 White Light and 21 Narrow Band Image) videos Categories were created from RGB images of these films at various polyp angles using VLC. 2700 photos per category were retrieved totaling 8100. We split the dataset into training and testing. The training set comprised 1890 photographs per class, whereas the testing set had 810. For validation, 1134 training photos were used. Careful dataset preparation supports research and model training [8]. Figure 2 illustrates model image pre-processing, Figure 3 the operation procedure.

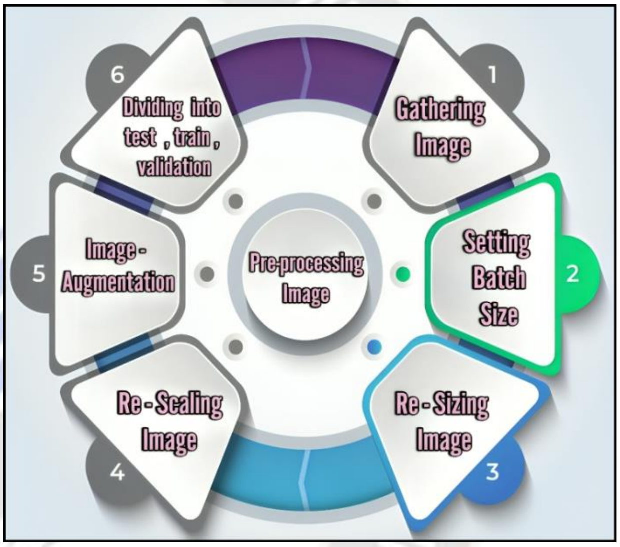


Figure 2 : Pre-processing of Image for proposed Model

IV. DATASET PREPROCESSING

The preparation of datasets for training and testing is a critical component of deep learning, machine learning, and data analysis. Nevertheless, this procedure is referred to as dataset pre-processing [9]. A raw dataset can contain a multitude of errors and undesirable values, which are commonly referred to as "noise". Deep learning systems that are pre-existing or algorithms that are custom-built would be incapable of training on such data. Conversely, data pre-processing resolves each inadequacy in the dataset. Furthermore, the model is capable of generating results that are satisfactory.

In the work we conducted, we implemented preprocessing techniques. An openly accessible polyp dataset that had previously undergone extensive cleansing and annotation served as the basis for our research. However, the dataset was initially in video format, necessitating its conversion to photos

International Journal on Recent and Innovation Trends in Computing and Communication
***ISSN:** 2321-8169 **Volume:** 13 **Issue:** 1*
***Article Received:** 25 November 2024 **Revised:** 12 December 2024 **Accepted:** 30 January 2025*

from various perspectives. It was subsequently discovered that the picture collection contained photographs of varying sizes, which required standardisation. Furthermore, in order to optimize the efficacy of the model we constructed and currently state-of-the-art (SOTA) models, it was necessary to resize and rescale the images, with a particular focus on the Karras API [10]. Concurrently, the effective training of deep learning (DL) models necessitates the use of large datasets. Additionally, image augmentation emerged as a key technique for data preparation prior to training convolutional neural network (CNN) models [11].

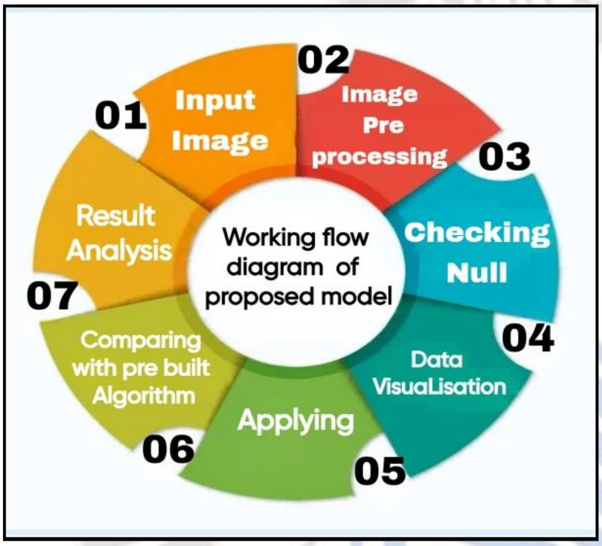


Figure 3: Working flow of the proposed model

A structured collection of 128 photographs with labels was produced by utilising TensorFlow's dataset API. Each image comprises three colour channels and 256x256 pixels. In order to improve the convergence of the model and the speed of the computations, we use a pretreatment pipeline inside we re-size and rescale sequential model to normalise the values of the pixels and standardise the dimensions of the image [12]. The aforementioned integration underscores the critical role they play in harmonising and refining picture data, thereby significantly contributing to the model's overall efficacy and resilience. Additionally, the designated resizing value of 256x256 pixels and the rescaling value of 1./255 ensure consistency in picture dimensions and the ease with which pixel intensities may be normalised. These values are essential for the efficient training of models and the adaptation of datasets [13][14].

Through the process of applying a variety of changes to the original picture, the process of image augmentation entails the creation of several changed copies of the same image. Nevertheless, each image reproduction is distinct from the other [15]. Consequently, the success of the procedure is contingent upon the augmentation techniques employed. Thanks to this procedure, the CNN model is able to more effectively generalise unknown data, which also increases the dataset's size and introduces a degree of variance.

This work employs the tf.keras Keras API. Layers is the sole layer in the sequential model known as data_augmentation Random Rotation (0.02). The maximal rotation angle of the images in the dataset is 0.02 radians or 1.15 degrees and they are subjected to random rotations by this layer.We effectively carried out image enhancement through a random rotation of up to 1.15 degrees and subsequent rescaling. To normalise the photos, the original 0–256 colour value range of each pixel is divided by 255, resulting in a new range of : (0–1.0). The aforementioned method is referred to as rescaling [16][17]

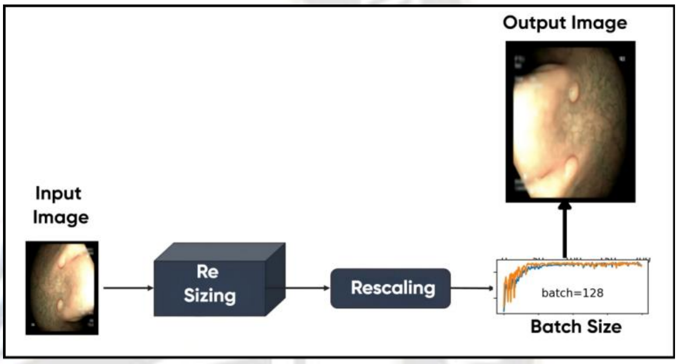


Figure 4: Re-sizing and Re-scaling on polyp Image

. Figure 4 illustrates the process of resizing and scaling a polyp image, while Figure 5 illustrates the classification of colon polyps using the sequential convolutional neural network (CNN) system that was recommended.

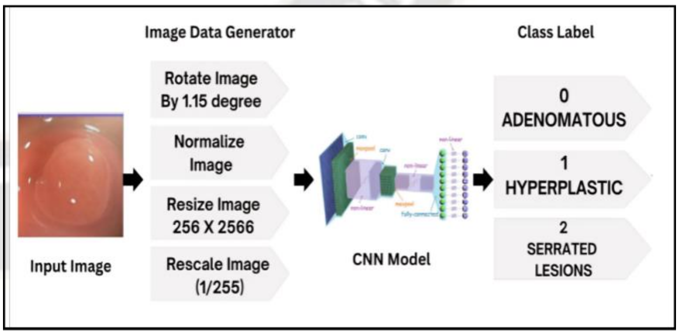


Figure 5: Overall Colon Polyps Classification Process with the Proposed sequential CNN Framework

V. CLASSIFICATION PROCESS

Our suggested study used a six-layer sequential convolutional neural network (CNN) model for polyp detection to address lengthy training periods and precise outcomes. Hardware environment matters in CNN model training for picture categorization. We used Kaggle's GPU T4x200 for computational aid without a dedicated GPU despite our willingness to persevere. I used my reliable workstation "DESKTOP-OF8J2FL" with an Intel(R) Pentium(R) CPU N4200 @ 1.10GHz and 8GB RAM for my research.

Specifically designed for applications that involve the classification of images, we suggest a convolutional neural network (CNN) model. Moreover, the architecture extends over six convolutional layers, each of which is equipped with dropout regularisation and max-pooling layers to mitigate overfitting. Besides, the input images are straight away scaled to a range of 0 to 2 and subsequently reduced to 256x256 pixels with RGB channels as part of the preparation phase. The convolutional layers are followed by the 2x2 max-pooling layers, which shrink the feature maps and include ReLU activation functions and 3x3 kernels. Dropout layers are intentionally incorporated after each convolutional layer to prevent overfitting. Consequently, 20% of the nodes are haphazardly eliminated by dropout layers during the training process. Following the contrary, the flattened feature maps are covered with two dense layers that are entirely connected. The initial dense layer, which is composed of 32 neurones and activated with ReLU, functions as a bottleneck for feature compression. In conclusion, the output layer employs softmax activation to construct probability distributions for each of the output classes in order to perform multi-class classification. Nevertheless, the Adam optimiser is employed to construct the model, which dynamically modifies the learning rate during training to optimise model parameters. Due to its superior performance in multi-class classification tasks, sparse categorical cross-entropy is implemented in loss computation. The accuracy measure is employed to evaluate the model's performance during training and assessment. Additionally, this section stipulates a comprehensive examination of the CNN model's architecture, preprocessing steps, standardising techniques and training setup. As a result, our model has a lower total number of parameters (175,427) than certain other models.

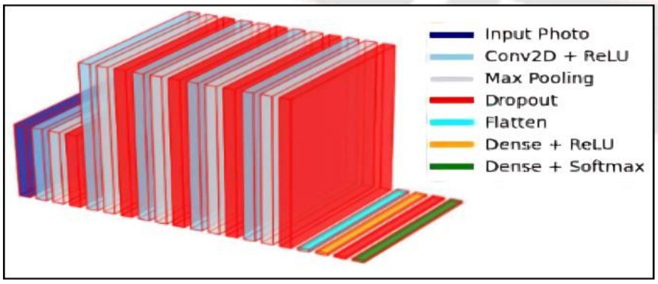


Figure 6: Proposed CNN Architecture

A 3D architectural block design of a convolutional neural network is depicted in figure 6, with each layer's coordinates, dimensions, labels, and colours. The architecture was simplified and layer categories, such as input, convolutional, pooling, dropout, flattening and dense tuples were described in the visualisation. In order to get the best outcome possible from the model, all of the parameters shown in "Table 1" are carefully adjusted. A total of 50 epochs were used to train the CNN model. Our proposed polyp classification CNN model's precise layer layout is listed in "Table 1".

The CNN model fitting technique is shown step-by-step in "Algorithm-1" with the pre-processing steps chosen based on the model type.

Algorithm-1 Proposed Colon Polyp Classification Algorithm

Load Dataset,

Then, install the pip of U-tensorflow,

Pre-processed Dataset,

Epoch assign 50,

Visualizing the batch of the dataset,

Model,

If** model = CNN **then,

Set batch_size where batch size is 128,

Resize_image_assign as 256,

Rotate Image , [By the help of Image Augmentation],

Rescale = 1/255,

There is no missing value,

***Return** Processed_Data,*

Fit the model,

***If the** output is 0,*

***Then** Adenomatous,*

***Else if** the output is 1,*

***Then** Hyperplastic,*

***Else if** the output is 2,*

***Then** Serrated Lesions,*

Exit the model,

***Then,** evaluate the model on the test Dataset,*

***Then**, plot the dataset,*

***Again**, plot the dataset,*

***So**, that the graph can be more smooth,*

***Then**, Generate classification report, confussion matrix, AUC score*

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
***Article Received:** 25 November 2024 **Revised:** 12 December 2024 **Accepted:** 30 January 2025*

Finish the process

VI. PERFORMANCE EVALUATION

For the purpose of determining whether or not the model is effective, we estimated the accuracy, F1-score and area under the curve (AUC) for one dataset. Additionally, we computed the F1-Score and AUC, which can provide comprehensive support for both true and deceptive values, as accuracy is insufficient to establish the algorithm's complete validation [18][19][20]. Additionally, we determined the average duration of performance.

VII. MODEL EVALUATION

We have compared our proposed CNN model with other deep convolutional neural networks, each with its own set of advantages and disadvantages, including VGG-16, VGG-19, MobileNet, DenseNet-201, DenseNet-121, Inception- ResNet-V2, Inception-V3, Xception, EfficientNet B7 and ResNet-152-V2[4][21]. They can achieve high accuracy with reduced computing costs, have a lightweight design, and are efficient in picture categorization. Furthermore, these networks are renowned for their simplicity and precision.

TABLE I

LAYER ARRANGEMENT AND TOTAL PARAMETERS OF THE MODEL

Layer Type	Layer Arrangement and Out Shape
Convolutional Layer 1	32 filters, kernel size (3,3), Activation Function ReLU (128, 254, 254, 32)
MaxPooling Layer 1	Pool size (2,2) (128, 127, 127, 32)
Dropout Layer 1	Rate 0.2 (128, 127, 127, 32)
Convolutional Layer 2	64 filters, kernel size (3,3), Activation Function ReLU (128, 125, 125, 64)
MaxPooling Layer 2	Pool size (2,2) (128, 62, 62, 64)
Dropout Layer 2	Rate 0.2 (128, 62, 62, 64)
Convolutional Layer 3	64 filters, kernel size (3,3), Activation Function ReLU (128, 60, 60, 64)
MaxPooling Layer 3	Pool size (2,2) (128, 30, 30, 64)
Dropout Layer 3	Rate 0.2 (128, 30, 30, 64)
Convolutional Layer 4	64 filters, kernel size (3,3), Activation Function ReLU (128, 28, 28, 64)
MaxPooling Layer 4	Pool size (2,2) (128, 14, 14, 64)
Dropout Layer 4	Rate 0.2 (128, 14, 14, 64)
Convolutional Layer 5	64 filters, kernel size (3,3), Activation Function ReLU (128, 12, 12, 64)
MaxPooling Layer 5	Pool size (2,2) (128, 6, 6, 64)
Dropout Layer 5	Rate 0.2 (128, 6, 6, 64)
Convolutional Layer 6	64 filters, kernel size (3,3), Activation Function ReLU (128, 4, 4, 64)
MaxPooling Layer 6	Pool size (2,2) (128, 2, 2, 64)
Dropout Layer 6	Rate 0.2 (128, 2, 2, 64)
Flatten Layer	(128,256)
Dense Layer 1	32 neurons, Activation Function ReLU (128,32)
Dropout Layer 7	Rate 0.2 (128,32)
Dense Layer 2	3 neurons, Activation Function Softmax (128,3)

TABLE II

PERFORMANCE MEASUREMENT OF ALL THE SOTA MODELS WITH OUR MODEL

Method	**F1 Score**	**AUC**	**Accuracy**	**Average Training Time**

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
***Article Received:** 25 November 2024 **Revised:** 12 December 2024 **Accepted:** 30 January 2025*

VGG-16	98.67%	98.9%	97.55%	51 Min 24s
VGG-19	98.67%	98.9%	98.41%	45 Min 2s
Mobile-Net	1.00%	99.9%	99.85%	27 Min 35s
Dense-Net 201	94.00%	98.9%	94.35%	47 Min 19s
Dense-Net 121	98.67%	99.9%	98.55%	36 Min 54s
Inception -V3	95.33%	99.36 %	95.62%	33 Min 36s
ResNet-152-V2	94.33%	96.96 %	94.56%	60 Min 39s
Inception Resnet-V2	74.33%	89.22 %	75.30%	52 Min 1s
Xception	95.66%	99.92 %	96.77%	37 Min 21s
EfficientNet B7	98.66%	99.59 %	98.53%	74 Min 15s
Proposed Model	**99.67%**	**99.67 %**	**99.95%**	**18 Min 59S**

VIII. EXPERIMENTAL RESULT

The outcomes of this investigation are substantial when our proposed CNN model is implemented to identify a variety of polyp categories. Nevertheless, deep learning techniques have also been effective. Additionally, our primary goal is to develop a model that can classify polyps with minimal training. Our proposed model comprises six convolutional layers, six "max pooling" layers to facilitate the extraction of low-level information, seven "dropouts" that safeguard the model from overfitting and two "dense layers" that establish a fully connected network by connecting each neurone in one layer to the neurone in the other layer. Furthermore, it is capable of transforming the input data into a format that can be employed to address particular issues, such as classification or regression. Bizarrely, this leads to a total of 175,427 parameters, which considerably contributes to the brief training period. Despite the fact that other per-trained models, such as "VGG16", "VGG19", "Mobile Net", "Dense Net-201", "Dense Net-121", "Inception ResNet V2", "Inception-V3", "Xception", "EfficientNet B7" and "ResNet 152-V2" result in larger parameters, the following are the largest: 14,812,995, 20,099,651, 3,425,475, 18,690,627, 7,234,115, 54,502,627, 22,023,971, 20,861,480, 64,589,210, and 58,724,867. In order to train the CNN model and the pertrained deep learning model, a total of fifty epochs are used. Our recommended model and all per-trained SOTA models are enumerated in **"Table-2"** along with their F-1 scores, AUC, accuracy and average training duration. Among the pretrained models, the proposed model achieved the highest accuracy of **99.95%** after **18 minutes and 59 seconds** of training. MobileNet, EfficientNet B7, DenseNet-121 and VGG-19 have exactly comparable accuracy of 99.85%, 98.53%, 98.55% and 98.41% to our model. Inception ResNet V2 has mediocre accuracy (75.30%). The following algorithms performed poorly: Xception (96.77%), DenseNet-201 (94.35%), InceptionV3 (95.62%), ResNet-152V2 (94.56%), and VGG-16 (97.50%). We came to realize that our model trains swiftly and most accurately. Figure 7 presents a bar chart of Table 2 variables F1 score, AUC, accuracy, and average training time for comparison.

Our proposed model outperforms pre-trained State-of-the-Art (SOTA) models in F1 score, AUC, accuracy and average training time as demonstrated by Table 2 and Figure 6.The majority of the models demonstrated exceptional performance, with scores exceeding 90%, with the exception of Inception ResNet V2, achieved only 75.30% accuracy. Nevertheless, this implies that our model offers evident advantages in terms of efficiency and accuracy, even in the face of success from other models. Moreover, it is imperative to evaluate the accuracy of the model training and validation, as well as the loss of the model during training and validation, in the context of deep learning [22][23]. These graphs demonstrate the extent to which a model learns over time. Fig 8 illustrates these performance graphs, which contrast our proposed model with other pre-trained models. (a) Our Proposed Model, (b) VGG-16, (c) VGG-19, (d) DenseNet-201, (e) Xception, (f) MobileNet, (g) EfficientNet B7, (h)DenseNet-121, (i) Inception ResNet V2, (j) Inception V3, and (k) ResNet-152 V2.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
***Article Received:** 25 November 2024 **Revised:** 12 December 2024 **Accepted:** 30 January 2025*

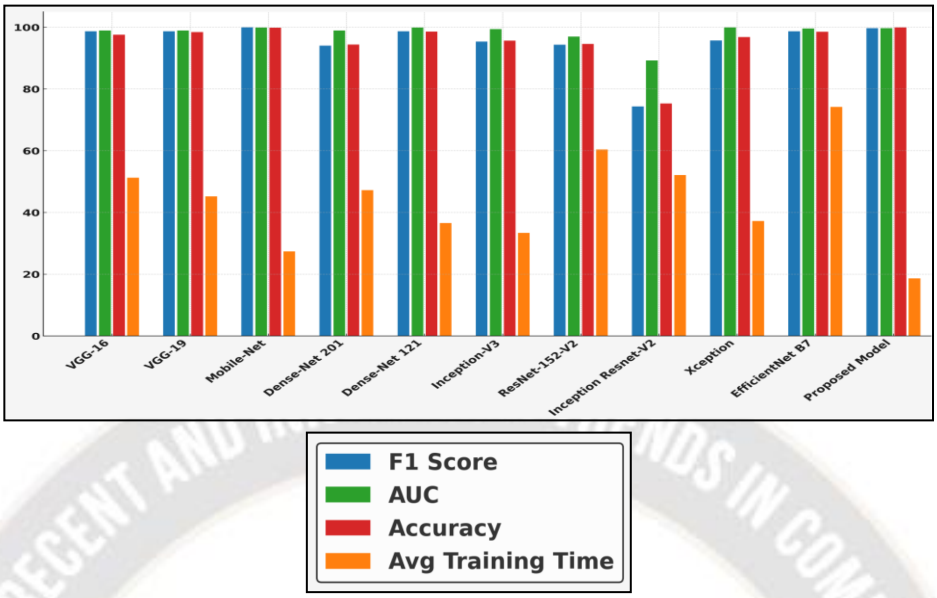


Figure 7: Performance Comparison of all the trained models

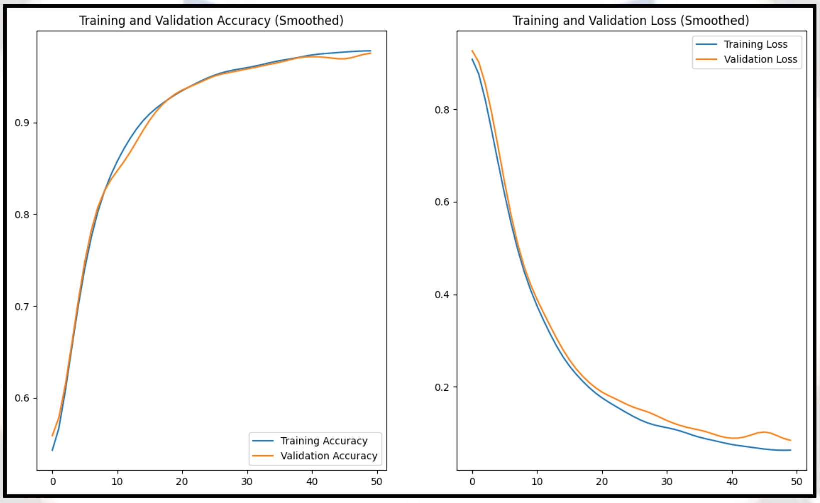


Figure 8.1: (a) Our Proposed Model

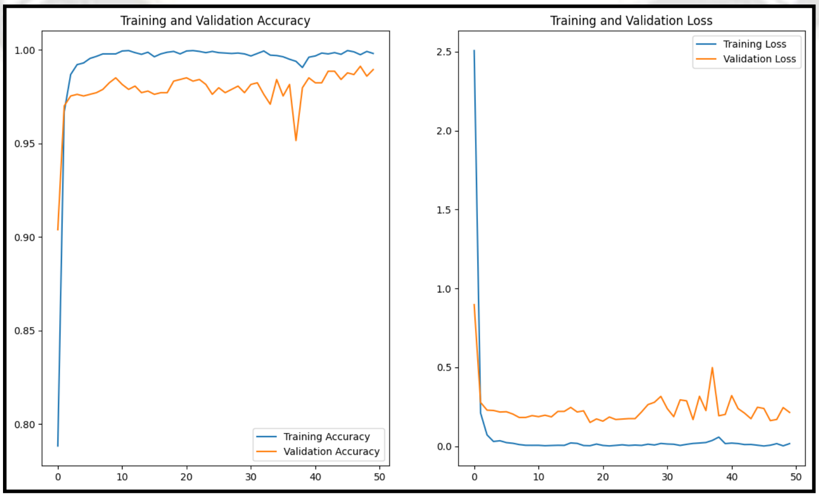


Figure 8.2: (b) VGG16

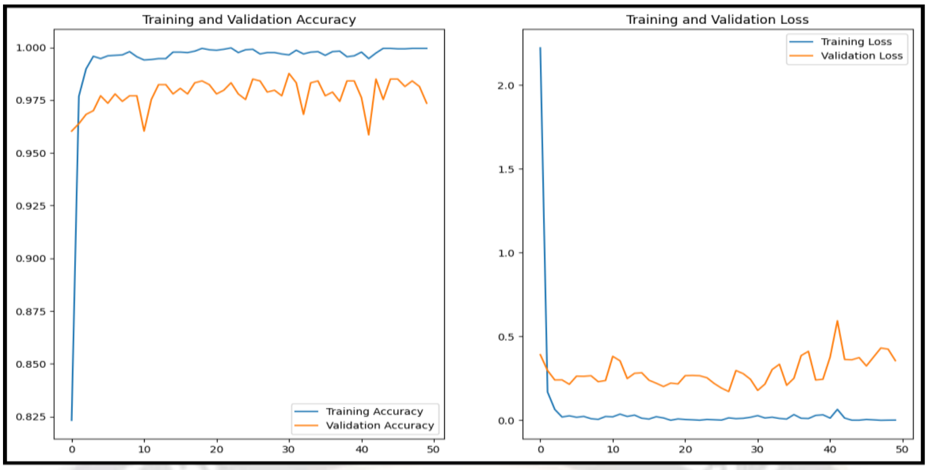


Figure 8.3: (c) VGG19

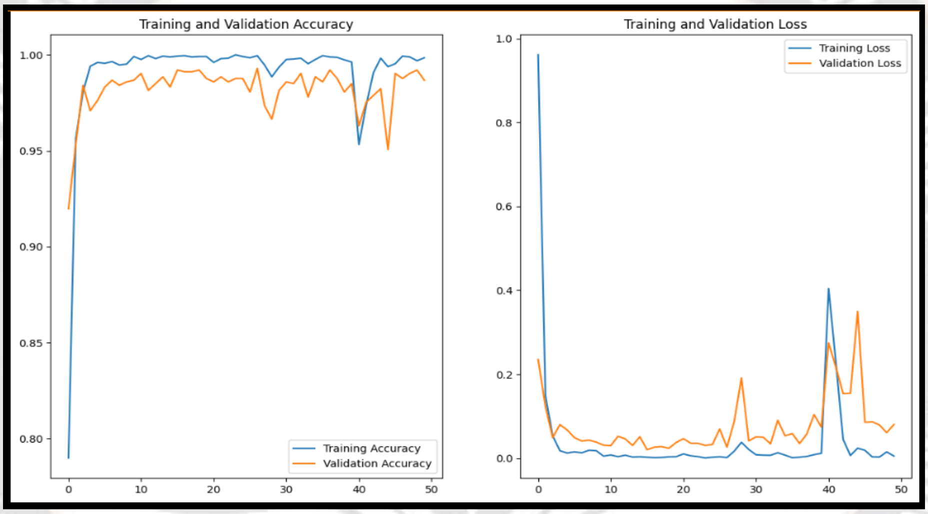


Figure 8.4: (d) DenseNet-201

Figure 8.5: (e) Xception

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
***Article Received:** 25 November 2024 **Revised:** 12 December 2024 **Accepted:** 30 January 2025*

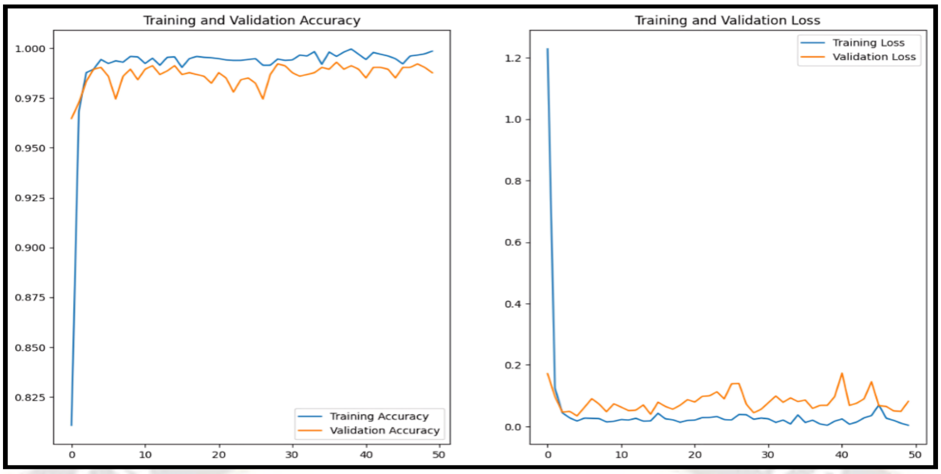


Figure 8.6: (f) MobileNet

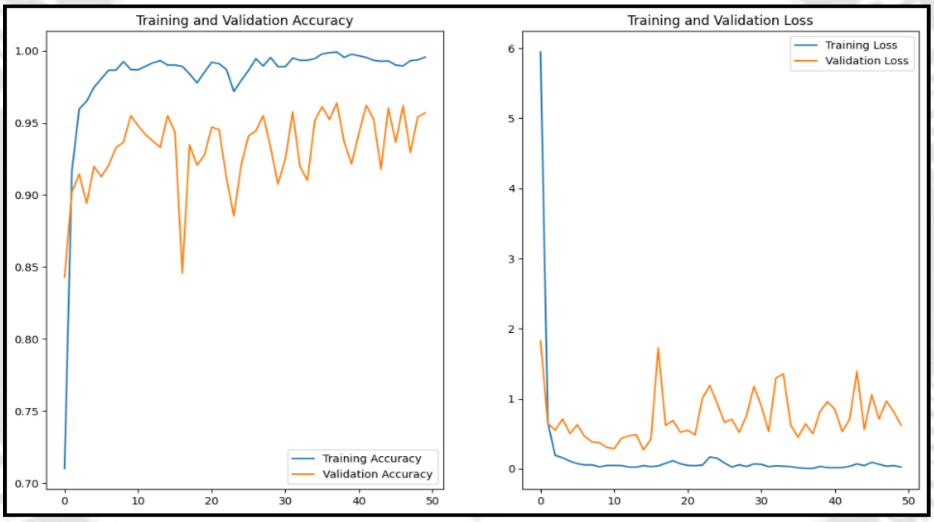


Figure 8.7: (g) EfficientNet B7

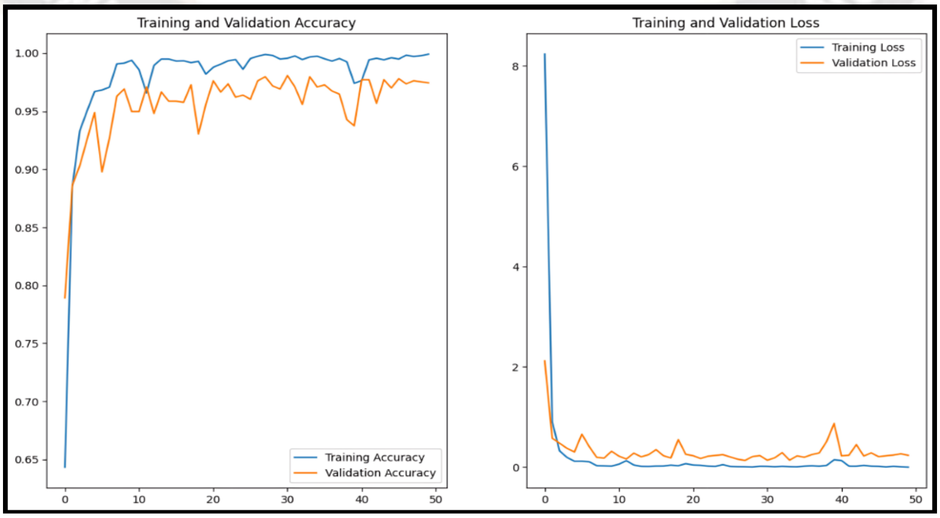


Figure 8.8: (h) DenseNet-121

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
***Article Received:** 25 November 2024 **Revised:** 12 December 2024 **Accepted:** 30 January 2025*

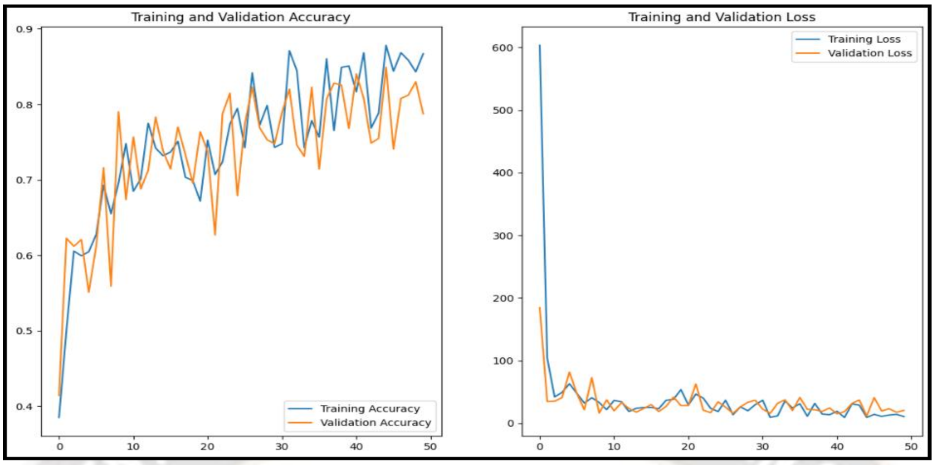


Figure 8.9: (i) Inception ResNet V2

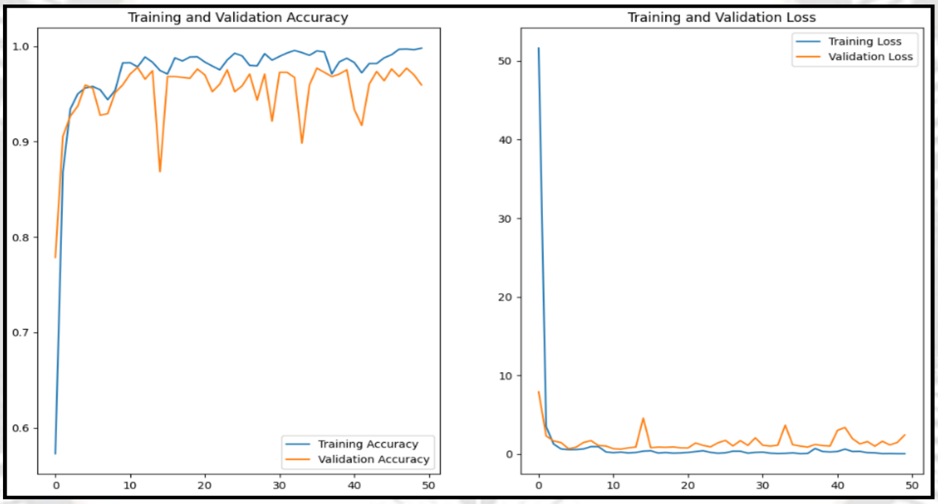


Figure 8.10: (j) Inception V3

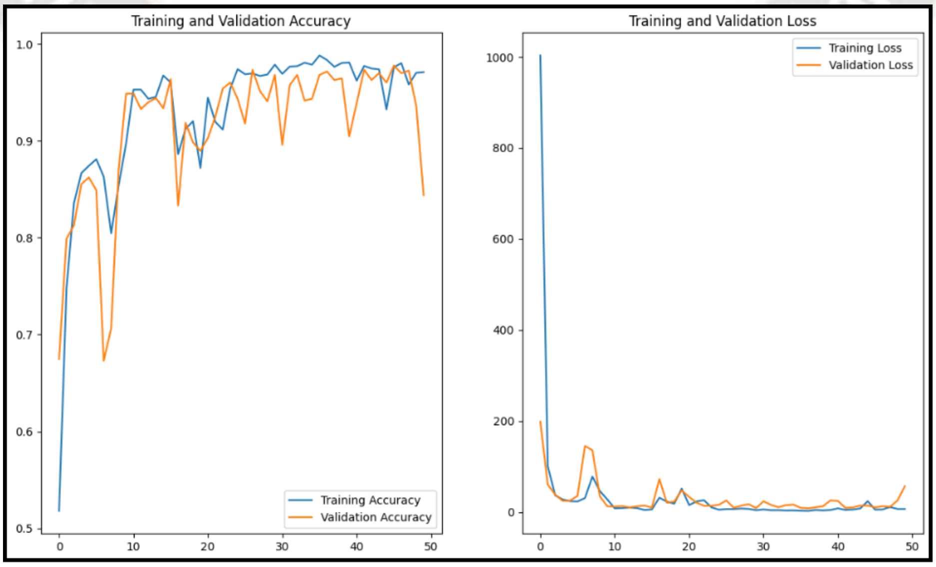


Figure 8.11: (k) ResNet-152 V2

The graph of our proposed model is the only one in Figure 8.1 that displays a uniform and consistent learning curve, which suggests that the model is capable of effective generalization and stable training. Other pre-trained models' graphs, on the other hand, exhibit anomalies like fluctuations or problems with convergence, which point to less-than-ideal learning behavior. Despite previous models' impressive accuracy, this one shows that their training procedures were inefficient and may have suffered from overfitting or underfitting [23]. Our suggested model's better learning curve shows how resilient it is and how well it can identify significant patterns in the data .

In comparison to alternative methods, Table 3 plainly demonstrates that our proposed model is quite acceptable and performs better. Our model's efficient and lightweight design is significantly enhanced by the reduced number of parameters in comparison to the model developed by Imrul Kayas [1]. Additionally, the results are devoid of transparency and reproducibility, as numerous authors fail to provide comprehensive descriptions of their models or divulge the code that implements them [24][25][26].

Conversely, our investigation offers a comprehensive account of the model architecture, training procedures, and performance indicators. Our proposed model exhibits exceptional performance in terms of accuracy, average training time, AUC, and F1 score. Meanwhile, the smoothness and excellent convergence of the learning curve demonstrate efficient learning and generalisation. The stability, lucidity, and usability of our model for real-world applications are enhanced by the collaborative efforts of all of these components.

TABLE III

EARLIER RELATED WORKS V/S OUR PRESENT WORK

Paper Author & Reference	Model Name	Performance
Md. Imrul Kayes [1]	Lightweighted CNN	97.6%
Farah Younas [2]	Compares six CNN model	96.3% , 81.2%
Pallabi Sharma [3]	LPNet	93.5%
Sushama Tanwar [4]	Shot MultiBox Detector (SSD)	92%
Win Sheng Liew [5]	use AI-based CAD tools	98.13%
Kangkana Bora [7]	NSCT	93.33%
Shin, Y.[8]	Combined feature +SVM,CNN(gray), CNN(RGB)	84.16%, 59.65%, 97.71%
Our Proposed Model	**Proposed Model**	**99.95%**

IX. CONCLUSION AND FUTURE WORKS

It is imperative to promptly identify polyps in order to facilitate the rapid detection and treatment of colon cancer, which is still one of the most prevalent and lethal malignancies worldwide. Additionally, this investigation demonstrated that a layered Convolutional Neural Network (CNN) model is capable of detecting polyps with an atypical accuracy of **99.95%**, despite the brevity of the **18 minutes and 59 seconds** of training.Compared to pre-trained State-of the-Art (SOTA) models, our data augmentation techniques significantly accelerated the model's setup and fine-tuned its performance. In order to enhance our CNN model, we intend to investigate alternative network designs and parameter sets in the future. Furthermore, our goal is to evaluate alternative methodologies in order to strengthen their adaptability and functionality. Beyond that, future research will focus on improvement of the generalisability and robustness of our methodology by adapting it to numerous local datasets, as this study was conducted using a singular dataset. On top of that, our ultimate objective is to develop a user-friendly application facilitating individuals to independently identify polyps and a comprehensive system for the automated diagnosis of early-stage colon nodules.

X. REFERENCES

[1] .Kayes, M. I., Prome, R. F., Noor, M., Bhowmik, S., & Ahmed, M. (2022, February). An Efficient and Lightweight Convolutional Neural Network for Carcinogenic Polyp Identification. In *2022 International Conference on Innovations in Science, Engineering and Technology (ICISET)* (pp. 278-283). IEEE.

[2] Younas, F., Usman, M., & Yan, W. Q. (2023). A deep ensemble learning method for colorectal polyp classification with optimized network parameters. *Applied Intelligence*, *53*(2), 2410-2433.

[3] Sharma, P., Das, D., Gautam, A., & Balabantaray, B. K. (2023). LPNet: A lightweight CNN with discrete wavelet pooling strategies for colon polyps classification. *International Journal of Imaging Systems and Technology*, *33*(2), 495-510.

[4] Tanwar, S., Vijayalakshmi, S., Sabharwal, M., Kaur, M., AlZubi, A. A., & Lee, H. N. (2022). Detection and classification of colorectal polyp using deep learning. *BioMed Research International*, *2022*.

[5] Liew, W. S., Tang, T. B., Lin, C. H., & Lu, C. K. (2021). Automatic colonic polyp detection using integration of modified deep residual convolutional neural network and ensemble learning approaches. *Computer Methods and Programs in Biomedicine*, *206*, 106114.

[6] Jha, D., Ali, S., Tomar, N. K., Johansen, H. D., Johansen, D., Rittscher, J., ... & Halvorsen, P. (2021). Real-time polyp detection, localization and segmentation in colonoscopy using deep learning. *Ieee Access*, *9*, 40496-40510.

[7] Bora, K., Bhuyan, M. K., Kasugai, K., Mallik, S., & Zhao, Z. (2021). Computational learning of features for automated colonic polyp classification. *Scientific Reports*, *11*(1), 4347.

[8] Shin, Younghak, and Ilangko Balasingham. ”Comparison of handcraft feature based SVM and CNN based deep learning framework for automatic polyp classification.” In 2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp. 3277-3280. IEEE, 2017

[9] Yamashita, Rikiya, Mizuho Nishio, Richard Kinh Gian Do, and Kaori Togashi. ”Convolutional neural networks: an overview and application in radiology.” Insights into imaging 9, no. 4 (2018): 611-62

[10] Ahamed, M. F., Syfullah, M. K., Sarkar, O., Islam, M. T., Nahiduzzaman, M., Islam, M. R., ... & Chowdhury, M. E. (2023). Irv2-net: A deep learning framework for enhanced polyp segmentation performance integrating inceptionresnetv2 and unet architecture with test time augmentation techniques. *Sensors*, *23*(18), 7724.

[11] Bhandari, A. (2020). Image Augmentation on the fly using Keras ImageDataGenerator. *Analytics Vidya*.

[12] Danon, D., Arar, M., Cohen-Or, D., & Shamir, A. (2021). Image resizing by reconstruction from deep features. *Computational Visual Media*, *7*(4), 453-466.

[13] Saponara, S., & Elhanashi, A. (2021, September). Impact of image resizing on deep learning detectors for training time and model performance. In *International Conference on Applications in Electronics Pervading Industry, Environment and Society* (pp. 10-17). Cham: Springer International Publishing.

[14] Sabottke, C. F., & Spieler, B. M. (2020). The effect of image resolution on deep learning in radiography. *Radiology: Artificial Intelligence*, *2*(1), e190015.

[15] Sengupta, P., Mehta, A., & Rana, P. S. (2023, July). Enhancing performance of deep learning models with a novel data augmentation approach. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.

[16] Liu, C., & Kirchner, M. (2019, July). CNN-based rescaling factor estimation. In *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security* (pp. 119-124).

[17] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*, *44*(7), 3523-3542.

[18] Yang, T., & Ying, Y. (2022). AUC maximization in the era of big data and AI: A survey. *ACM Computing Surveys*, *55*(8), 1-37.

[19] Vakili, M., Ghamsari, M., & Rezaei, M. (2020). Performance analysis and comparison of machine and deep learning algorithms for IoT data classification. *arXiv preprint arXiv:2001.09636*.

[20] Sulam, J., Ben-Ari, R., & Kisilev, P. (2017, September). Maximizing AUC with Deep Learning for Classification of Imbalanced Mammogram Datasets. In *VCBM* (pp. 131-135).

[21] Jeong, S. M., Lee, S. G., Seok, C. L., Lee, E. C., & Lee, J. Y. (2023). Lightweight Deep Learning Model for Real-Time Colorectal Polyp Segmentation. *Electronics*, *12*(9), 1962.

[22] Joodi, M. A., Saleh, M. H., & Kadhim, D. J. (2023). Increasing validation accuracy of a face mask detection by new deep learning model-based classification. *Indonesian Journal of Electrical Engineering and Computer Science*, *29*(1), 304-314.

[23] Chugh, G., Kumar, S., & Singh, N. (2024). TransNet: a comparative study on breast carcinoma diagnosis with classical machine learning and transfer learning paradigm. *Multimedia Tools and Applications*, *83*(11), 33855-33877.

[24] Jha, D., Smedsrud, P. H., Johansen, D., De Lange, T., Johansen, H. D., Halvorsen, P., & Riegler, M. A. (2021). A comprehensive study on colorectal polyp segmentation with ResUNet++, conditional random field and test-time augmentation. *IEEE journal of biomedical and health informatics*, *25*(6), 2029-2040.

[25] Shin, Y., Qadir, H. A., Aabakken, L., Bergsland, J., & Balasingham, I. (2018). Automatic colon polyp detection using region based deep CNN and post learning approaches. *IEEE Access*, *6*, 40950-40962.

[26] Zhang, R., Zheng, Y., Mak, T. W. C., Yu, R., Wong, S. H., Lau, J. Y., & Poon, C. C. (2016). Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain. *IEEE journal of biomedical and health informatics*, *21*(1), 41-47.

[27] Ribeiro, E., Uhl, A., Wimmer, G., & Häfner, M. (2016). Exploring deep learning and transfer learning for colonic polyp classification. *Computational and mathematical methods in medicine*, *2016*(1), 6584725.

[28] Hsu, C. M., Hsu, C. C., Hsu, Z. M., Shih, F. Y., Chang, M. L., & Chen, T. H. (2021). Colorectal polyp image detection and classification through grayscale images and deep learning. *Sensors*, *21*(18), 5995.

[29] 29. Hossain, M. S., Rahman, M. M., Syeed, M. M., Uddin, M. F., Hasan, M., Hossain, M. A., ... & Samad, M. A. (2023). DeepPoly: deep learning-based polyps segmentation and classification for autonomous colonoscopy examination. *IEEE Access*, *11*, 95889-95902.

[30] 30. Wang, W., Tian, J., Zhang, C., Luo, Y., Wang, X., & Li, J. (2020). An improved deep learning approach and its applications on colonic polyp images detection. *BMC Medical Imaging*, *20*, 1-14.