[1]Sushree Gayatri Priyadarsini Prusty
[2]Narayana Patra
[3]Adyasha Rath
[4]Debashish Swapnesh Kumar Nayak
[5]Jyotirmayee Rautaray
[1*]Sashikanta Prusty
[6]Nrusingha Tripathy
[6]Subrat Kumar Nayak

IntestiNet: Predicting Intestinal Abnormalities in Colon with Advanced CNN Techniques

Abstract-

Background: In previous years, colon cancer seems to be the second leading cause of death worldwide, which needs preventive strategies to combat this disease. Thus, timely diagnosis with effective methods may improve the survival rate of patients.

Method: This article presents an advanced method as IntestiNet for the classification and prediction of colon cancer which leverages the abilities of CNN models in three major steps: (i) train and validate the CNN model with 54% training data (i.e. 3200 images) and 13% validation data (i.e. 800 images), (ii) parameter optimization using Adamax and custom callbacks to continue or halt training process, and (iii) test CNN models with 33% data (i.e. 2000 images). We herein created train_gen, test_gen and valid_gen using ImageDataGenerator class with image size of (200, 250) which further passed into six CNN models (EfficientNetB2, ResNet50, InceptionV3, VGG16, MobileNetv2, and EfficientNetB5) for the classification of four types of classes as '0: Normal', '1: Ulcerative colitis', '2: Polyp', and '3: Esophagitis'.

Results: To validate the effectiveness of CNN models, an interactive approach using Keras applications for multiple hidden layer implementation has been conducted on a curated colon dataset. The results indicate how well the EfficientNetB5 model succeeds compared to other methods with an accuracy of 98.62%, precision of 98.62%, recall of 98.63%, and F1-score of 100%. Such outcomes demonstrate the possible benefit of EfficientNetB5 in enhancing colon cancer identification and diagnoses.

Conclusion: The proposed architecture in this article, demonstrates impressive propels in automatically identifying and classifying four types of colons. Comprehensive testing on a variety of metrics revealed EfficientNetB5's improved performance, which highlights the technology's potential to improve colon disease reconnaissance and prognosis.

Keywords- Artificial intelligence; CNN; Curated colon dataset; Adam Max optimizer

I. INTRODUCTION

Cancer disorders are among the world's biggest causes of death. It is projected that 16.4 million cancer-related deaths will occur worldwide in 2040 [1], with colorectal cancer (CRC) ranking third in terms of frequency [2]. Early detection is essential to reducing the death rate from cancer. To obtain good oncological outcomes, a unique technology that can be utilized to detect cancer regions during the preoperative diagnostic stage and to guarantee tumour-free resection margins following cancer removal would be revolutionary. It is the second most dangerous and third most frequent cancer in the world, accounting for 1.8 million new cases and 881,000 deaths globally in 2018 [3]. It is one of the main cancers. With a few notable exceptions, the prevalence of colorectal cancer (CRC) is rising worldwide, especially in emerging nations, where changing dietary and lifestyle choices are probably contributing to a rise in early-onset CRC [4].

To reduce the impact of CRC, which accounts for over 10% of cancer diagnoses and over 9% of cancer-related mortality globally, efficient screening and diagnostic techniques are required [5]. There is now strong evidence that the incidence and death of CRC have significantly decreased in advanced countries as a result of population screening [6, 7]. The most effective method for detecting and preventing colorectal cancer (CRC) is colonoscopy; nevertheless, its effectiveness depends on the expertise and attentiveness of the endoscopist [8]. Despite improvements in adenoma identification and endoscopic technology, adenoma miss rates following colonoscopy investigations can still reach up to 26% [9].

[1, 5, 6]Reserch Scholar, Computer Science & Engineering, Siksha O Anusandhan University, Bhubaneswar, India.
[2]Associate Professor, Computer Science & Information Technology, Siksha O Anusandhan University, Bhubaneswar, India.
[3]Assistant Professor, Computer Science and Engineering, C. V. Raman Global University, Bhubaneswar, India
[4]Assistant Professor, Computer Science and Engineering, Centurion University of Technology & Management, Bhubaneswar, India
liza.sushree19@gmail.com, narayanpatra@soa.ac.in, adyasha.rath@cgu-odisha.ac.in, swapnesh.nayak@gmail.com jyotirmayee.1990@gmail.com, nrusinghatripathy654@gmail.com, subratsilicon28@gmail.com, sashi.prusty79@gmail.com*
Correspondence Author: mailto: sashi.prusty79@gmail.com

Artificial intelligence (AI) integration has become a revolutionary force in modern healthcare, revolutionizing many facets of medical practice [10-13]. Numerous fields of medical science have investigated AI, with encouraging outcomes. By supplementing human expertise and possibly lowering the miss rates of these subtle polyps, the inclusion of AI into colonoscopy holds promise for improving patient outcomes. Davri et al. [14] offer a comprehensive assessment of the use of DL in colorectal cancer using computerized image analysis on histopathology images. Furthermore, the constraints were described to motivate researchers to offer solutions. Dalai et al. [15] examined research that used machine learning (ML) methods to predict colorectal cancer. When it comes to cancer prediction, ML has shown to be incredibly effective.

The key contributions in this research are described as follows:

- We proposed a novel classification and prediction method using six different CNN models on a large curated colon cancer dataset (i.e. 6000 normal and abnormal images). This method classifies the classes based on the train and validation set, whereas predicts the appropriate class based on the test set.
- We introduce the trim () function to reduce the training time limit to the maximum and to allocate 500 samples for each four classes.
- We herein implemented the ImageDataGenerator class for augmenting using train_gen, valid_gen, and test_gen objects on batch_size =30.
- We present an optimizing parameter and initiate Keras callbacks block that uses Adamax optimizer at learning rate=0.001 and 2 necessary callbacks to handle learning rate and early stop.
- We perform the polyp prediction on 33% curated test dataset (i.e. 800 samples) using CNN models.
- To determine the efficacy and adaptability of the six algorithms, it is essential to analyze their performance on an increased image dataset.
- To demonstrate the resilience and flexibility of our approach, we assess its performance using six models where EfficientNetB5 emerges as the top predictive model.

The rest of this paper includes Section 2, which describes the proposed method architecture to predict the intestine abnormalities in the colon; Section 3, which discusses the results obtained; and Section 4 concludes with a conclusion and future perspectives.

II. RESEARCH METHODS

This research article offers a proposed method of architecture, which has been developed specifically for the accurate classification and prediction of four important classes in meticulously chosen colon images as in Figure 1. The main goal of this method is to quickly detect and distinguish polyps, which are important markers for many gastrointestinal disorders. The encoder of the proposed method comprises 3 basic steps (i) import needy modules, (ii) experimental setup, and (iii) CNN models implementation. The last step further classifies into three basic steps i.e. training CNN models with 2000 samples, optimizing parameters and initiating callbacks, and predicting CNN models with 800 test samples.

A. Import Needy Modules

The images in the digital world have pixels whose intensities range from 0 to 255 where black is 0 and white is 255. The rescaling of 1./255 allows it to convert to decimal values ranging from 0 to the flow_from_directory allowing us to use the images available in the Google Colab files section. The dataset contains images of variable sizes. Hence it is essential to reshape them (224,224). The batch size is an important parameter that can affect the training time and accuracy. Less batch size increases accuracy and training time, so the batch size is taken as powers of two (i.e. here batch_size=30).

B. Experimental Setup

A Lenovo Z50-70 Intel(R) Core(TM) i7-4500U CPU @ 2.40GHz, 8 GB RAM, and NVIDIA® GeForce® 840M connectivity were used for the CNN model training and testing. The Jupytor Notebook platform and the Python programming language were used to experiment. The dataset starts with 6000 raw photos, and trim () on

the train set is used to further reduce the training time. This yields about 2,000 photos, split evenly between the four target classes. Using 12 epochs, this training procedure takes about 3 to 6 hours.

C. CNN Models Implementation

We are using the pretrained CNN meaning that the CNN is already trained on the imagenet dataset and has the weights stored in it. The input dimension for CNN models is to be input_shape = (224, 224, 3). We are training it partially to reduce the training time on the stake of accuracy.

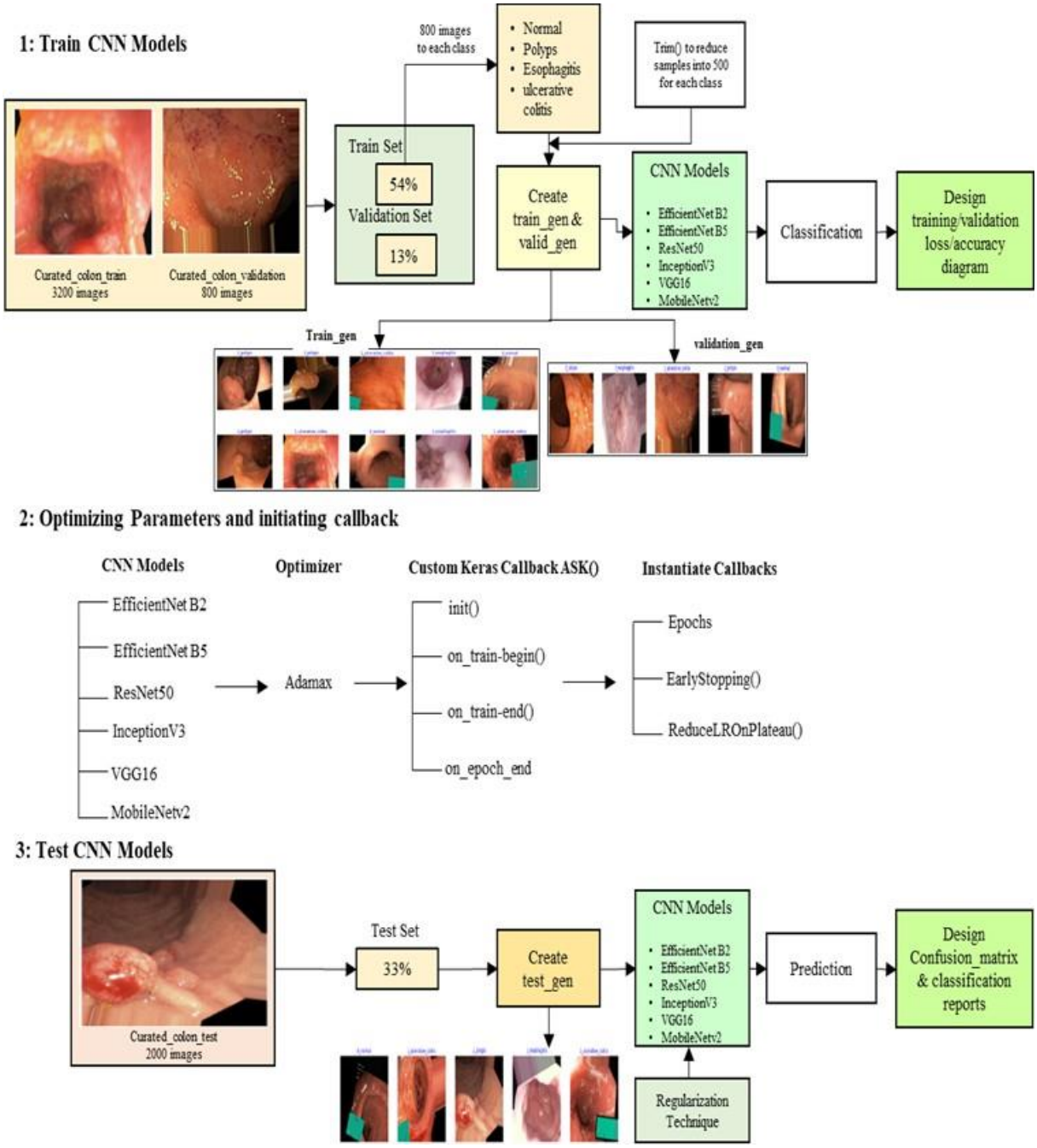

Fig. 1. Proposed IntestiNet method design for the classification and prediction of intestine abnormalities if present in curated colon dataset

Algorithm 1:

Input: ***curated-colon-dataset***

1. Import basic nodules
2. Data preprocessing
   a) *Split train set, validation set, and test set*
   b) *train_df length: 3200, test_df length: 800 valid_df length: 2000*
   c) *equally distributed images into four classes 0_normal, 1_ulcerative_colitis, 2_polyps, 3_esophagitis*
   d) *trim train_df to reduce time*
   e) *allow a maximum of 500 samples to each of four classes i.e. [500, 500, 500, 500]*
   f) *Augmenting with ImageDataGenerator class and create train_gen, test_gen, & valid_gen.*
3. Train CNN models at learning rate = 0.001
4. Optimizing with Adamax parameter and initiating custom callbacks to control learning rate & early stopping
5. Plot training and validation data w.r.t 12 epochs
6. Predict 800 test samples
7. Design confusion matrix and classification results
8. Save the model

1) Train CNN Models: In the beginning stage, we take a curated colon image dataset consisting of 3200 train (i.e. 54%) and 800 (i.e. 13%) validation images with a 150*150 pixel format. This train set consists of 800 images for each four classes as shown in Figure 1. Since the images are more complex to analyze, we herein trim the training set by reducing the number of samples to 500 for each class. Further, augmenting images using the ImageDataGenerator () class using train_gen and valid_gen objects (Figure 2).

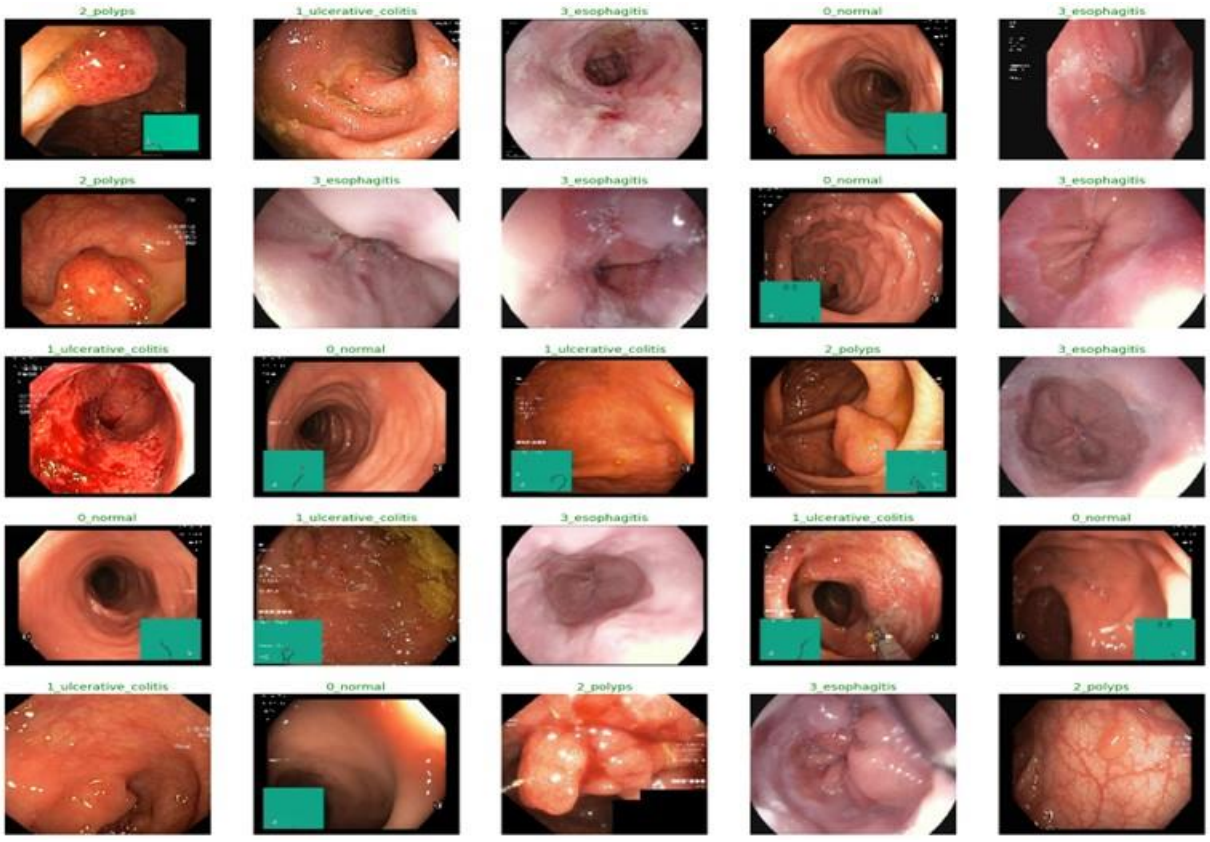

Fig. 2. Identifying four target class classifications using the train_gen object of the ImageDataGenerator () class

This takes less memory and makes sure the model gets unique image variations at every epoch. The number of image batches in a single epoch is specified by steps_per_epoch, which is usually calculated by dividing the

dataset length by batch size (i.e. here is 30). The outcome shows 2000 images significantly distributed into 4 classes. CNN models further proceed with those validated images passing through hidden layers for the model compilation at learning_rate = 0.001 and loss = 'categorical_crossentropy'. During this phase, the 'relu' activation function ignores the negative inputs and produces the exact match for positive values.

2) Optimizing Parameters & Initiating Callbacks: In this study, we have used the Adamax optimization technique, an extension of Adam used to accelerate the optimization process. At first, to optimize any of the parameters throughout the search, let's say M=0 and U=0, respectively, we need to maintain a moment vector and an exponentially weighted infinity norm.

Beginning at t=1, the procedure is run iteratively over time t, calculating a new set of parameter values p at each iteration $s.t.\ p(t-1) \dots\dots\ p(t)$. The gradient (partial derivatives) for the current time step are first computed.

$$gr(t) = f'(p(t-1)) \tag{1}$$

Next, we tried with the gradient and a hyperparameter $h1$ to update the moment vector $M(t)$ as:

$$M(t) = (1-h1) * gr(t) + h1 * M(t-1) \tag{2}$$

Furthermore, we have taken hyperparameter $h2$, to update the exponentially weighted infinity norm $U(t)$ as:

$$U(t) = abs\big(gr(t)\big) + \max\,(h2 * U(t-1)) \tag{3}$$

Whereas, $abs()$ determines the absolute value and max() chooses the maximum of the parameters.

Deep learning model training is an extremely difficult procedure. It is nearly difficult to predict so many model parameters when the model is being trained. In Keras, a callback is performed when called using Keras. callbacks.Callback class that is used to view internal states and statistics of a model during the training phase. It is an object that we can call at various points during the training process and pass to the model while utilizing the fit method. At the point where the validation loss is no longer improving, the ReduceLROnPlateau () class is used to lower the learning rate and is particularly useful for emerging out of the local minima. When the validation loss is no longer improving, EarlyStopping () has been used to halt the training process. To guarantee that an optimal model is being stored throughout the training process, each model's weight has been maintained at various stages of training. In artificial intelligence (AI), epochs are crucial because they allow the model to learn and modify its parameters, resulting in better performance and generalization. During the model training process, choosing the right number of epochs is essential to striking the best possible balance between efficiency and learning accuracy.

3) Test CNN Models: Classifying the histology images using each CNN model's neural network architecture required training the model to recognize the types from various tissue classes. This was the final step in our architectural parameters. The curated colon dataset was utilized for testing after neural networks were trained on all 2,000 image patches (Figure 3). The classification metrics such as AUC (area under the curve), ROC (receiver operating characteristic) curve, accuracy, precision, recall, and F1-score used to quantify each model's performance on four target class classifications (i.e. 0, 1, 2, and 3) that were expanded from their binary equivalent. The accuracy is calculated as the ratio of correct predictions over all guesses (blue rectangle box as in Figure 4), based on a four-class confusion matrix (Figure 5). Using mathematical formulas, the category-wise matrices (recall, F1-score, and precision) were determined as in eqs. 4 to 8. Using 10 epochs and

800 test data, the EfficientNetB5 achieved the best accuracy of 98.62% than the other five CNN models as shown in Table 1 in this study.

III. RESULT & DISCUSSION

A. CNN models train, validation, and test scores

This study comprises colon cancer classification using six CNN models (i.e. EfficientNetB2, ResNet50, InceptionV3, VGG16, MobileNetv2, and EfficientNetB5) at 54% train set and 13% validation test; prediction of each model using 33% test set. Table 2 specifies each model's performances using the most common deep learning metrics (i.e. loss, accuracy, precision, recall, ROC curve, F1-score) on the train set, validation set, and test set. The performance comparison for all six models was trained on 2000 samples and tested on 800 samples for four class predictions. The results in Table 2 show that the EfficientNetB5 model achieves the highest accuracy and lowest loss scores in comparison to the other five models.

$$accuracy = \frac{tp+tn}{tp+tn+fp+fn} \quad (4)$$

$$precision = \frac{tp}{tp+fp} \quad (5)$$

$$recall = \frac{tp}{tp+fn} \quad (6)$$

$$F1-score = \frac{2*precision*recall}{precision+recall} \quad (7)$$

Table 1. Representation of each six CNN model's performance scores using classification_report class.

CNN models	Train/Test/Validation	Loss	Accuracy	Precision	Recall	ROC auc	F1-score
EfficientNetB2	train-set	0.0573	0.9806	0.9806	0.9806	0.9782	0.9716
	validation-set	0.2003	0.9265				
	test-set	0.1422	0.9463				
ResNet50	train-set	0.0817	0.99	0.97	0.97	0.96	0.97
	validation-set	0.0675	0.99				
	test-set	0.0749	0.98				
InceptionV3	train-set	0.0313	0.9884	0.9697	0.9636	0.9345	0.9699
	validation-set	0.2003	0.9265				
	test-set	0.2902	0.9325				
VGG16	train-set	0.0017	1.00	0.50	1.00	0.50	0.66
	validation-set	0.0048	0.9994				
	test-set	0.0364	0.9925				
MobileNetv2	train-set	0.0231	0.99	0.97	0.97	0.97	0.97
	validation-set	0.0715	0.96				
	test-set	0.0735	0.97				
EfficientNetB5	train-set	0.0323	0.9955	0.9862	0.9863	0.9862	1.0000
	validation-set	0.0308	0.9860				
	test-set	0.0049	0.9802				

B. EfficientNetB5 model performance diagram

The training result was obtained using 12 epochs on 2000 train image sets using different matrices such as loss vs. accuracy and val_loss vs. val_accuracy. The best scores were found at position epochs=12, as given in Figure 3. The training and validation accuracy were found nearly about 99% and loss to be less than 1%.

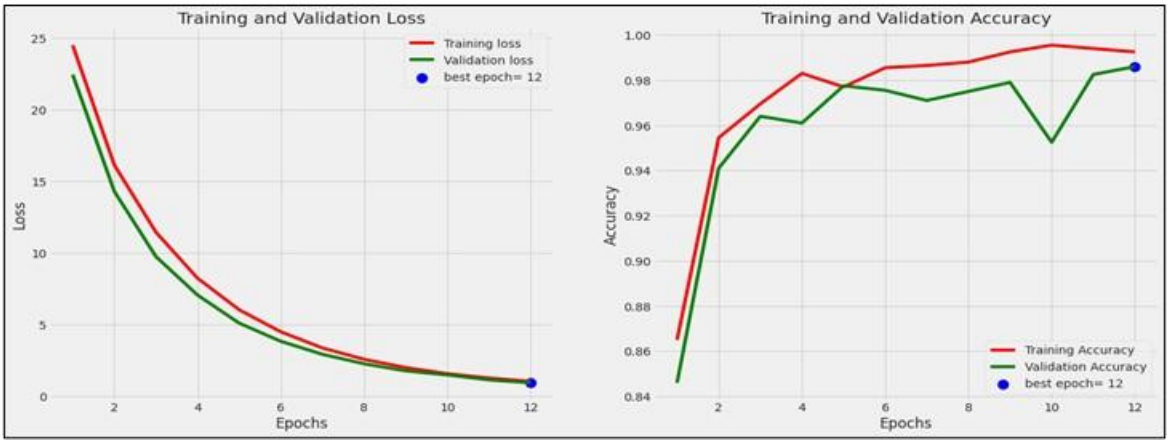

Fig. 3. Representation of training and validation loss/accuracy plot using 12 epochs on 2000 train image set

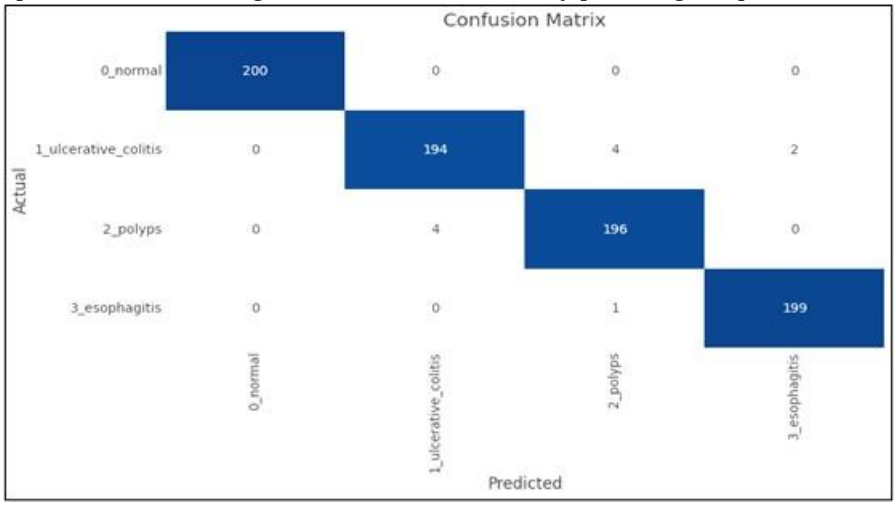

Fig. 4. Confusion matrix representation for the four classes in a blue rectangle box

C. Prediction using EfficientNetB5

We herein randomly take an image from the curated colon dataset to evaluate the appropriate class for that image. The image has checked for the respective class such as '0: Normal', '1: Ulcerative colitis', '2: Polyp', and '3: Esophagitis'. The model successfully predicts the class as '0: Normal' as shown in Figure 5.

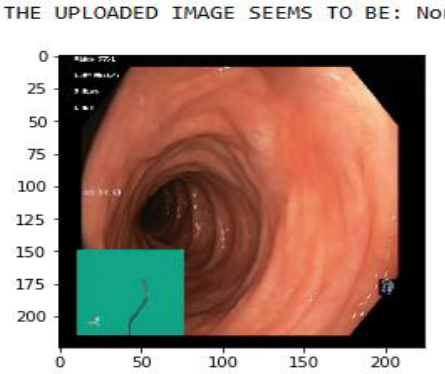

Fig. 5. EfficientNetB5 model prediction with four coordinates [[9.99474466e-01 4.56175534e-04 5.78074178e-05 1.15857865e-05]]

IV. CONCLUSION

This research paper presented a classification and prediction diagram of six different CNN models on a freely online available curated colon cancer dataset as in Figure 1 using three necessary steps. The result gathered as in Table 1, shows that the EfficientNetB5 CNN model experimentally proves as best model in terms of less loss and high accuracy percentages. Rather, other classification metrics such as precision, recall, ROC auc, and F1-score were calculated on the test set, where EfficientNetB5 proves as best concerning the other five models. The exceptional ability of the EfficientNetB5 to surpass cutting-edge techniques like ASK () and ImageDataGenerator () while employing a substantially large amount of images is one of its many interesting benefits. This diminishes the computing load and improves the model's efficiency in clinical settings where real-time data is exchanged. Finally, we go with the best-trained EfficientNetB5 model to predict the intestinal abnormalities if present in the input data using four class classifications. The model has successfully predicted the uploaded image as a normal class (Figure 5). In the future perspectives, we will evaluate multiple datasets and implement better integrative approaches to achieve better results.

Dataset Availability Statement

In this research, the experiment was done on a freely available curated colon dataset at https://www.kaggle.com/datasets/francismon/curated-colon-dataset-for-deep-learning.

REFERENCES

[1] Global Cancer Observatory. Available online: https://gco.iarc.fr (accessed on 29 March 2023).

[2] Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. 2018, 68, 394–424.

[3] Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. 2018, 68, 394–424.

[4] Siegel, R.L.; Torre, L.A.; Soerjomataram, I.; Hayes, R.B.; Bray, F.; Weber, T.K.; Jemal, A. Global patterns and trends in colorectal cancer incidence in young adults. Gut 2019, 68, 2179–2185.

[5] Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. 2018, 68, 394–424.

[6] Arnold, M.; Sierra, M.S.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global patterns and trends in colorectal cancer incidence and mortality. Gut 2017, 66, 683–691.

[7] Cardoso, R.; Guo, F.; Heisser, T.; Hackl, M.; Ihle, P.; De Schutter, H.; Van Damme, N.; Valerianova, Z.; Atanasov, T.; Májek, O.; et al. Colorectal cancer incidence, mortality, and stage distribution in European countries in the colorectal cancer screening era: An international population-based study. Lancet Oncol. 2021, 22, 1002–1013.

[8] Corley, D.A.; Jensen, C.D.; Marks, A.R.; Zhao, W.K.; Lee, J.K.; Doubeni, C.A.; Zauber, A.G.; de Boer, J.; Fireman, B.H.; Schottinger, J.E.; et al. Adenoma detection rate and risk of colorectal cancer and death. N. Engl. J. Med. 2014, 370, 1298–1306. [Google Scholar] [CrossRef]

[9] Zhao, S.; Wang, S.; Pan, P.; Xia, T.; Chang, X.; Yang, X.; Guo, L.; Meng, Q.; Yang, F.; Qian, W.; et al. Magnitude, Risk Factors, and Factors Associated with Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-analysis. Gastroenterology 2019, 156, 1661–1674.e11.

[10 Prusty, S., Patnaik, S., Dash, S. K., & Prusty, S. G. P. (2024). SEMeL-LR: An improvised modeling approach using a meta-learning algorithm to classify breast cancer. Engineering Applications of Artificial Intelligence, 129, 107630.

[11] Patil, P., Deokate, S. T., Bhoite, A., Prusty, S., Patil, A. B., & Mange, P. (2023). GAN-Enhanced Medical Image Synthesis: Augmenting CXR Data for Disease Diagnosis and Improving Deep Learning Performance. Journal of Electrical Systems, 19(3).

[12 Prusty, S., Das, P., Dash, S. K., Patnaik, S., & Prusty, S. G. P. (2023). Prediction of Breast cancer using integrated machine learning-fuzzy and dimension reduction techniques. Journal of Intelligent & Fuzzy Systems, 45(1), 1633-1652.

[13] Prusty, S., Patnaik, S., & Dash, S. K. (2022). SKCV: Stratified K-fold cross-validation on ML classifiers for predicting cervical cancer. Frontiers in Nanotechnology, 4, 972421.

[14] Davri, A., Birbas, E., Kanavos, T., Ntritsos, G., Giannakeas, N., Tzallas, A.T., Batistatou, A. Deep Learning on Histopathological Images for Colorectal Cancer Diagnosis: A Systematic Review. Diagnostics 2022, 12, 837.

[15] Dalai, S. S., Prusty, S., & Rautaray, J. (2024). Enhancing Transfer Learning Model Performances using Cohen's Kappa Score at Predicting Breast Cancer. International Journal of Intelligent Systems and Applications in Engineering, 12(11s), 290-303.