Undergraduate Conference on Intelligent Computing and Systems (UCICS 2025)
26-27 February, 2025; Varendra University, Rajshahi, Bangladesh

Automated Detection and Classification of Colon Cancer Using Transfer Learning with Deep CNN

Nitun Kumar Podder
Dept. of Computer Science & Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
E-mail: nituncse@gmail.com

Tania Sultana
Dept. of Computer Science & Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
E-mail: taniasultanaanni@gmail.com

Md. Sanzid Olioul Tanzid
Dept. of Computer Science & Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
E-mail: sanzid94@gmail.com

***Abstract*— Nowadays Cancer has become the most common cause of death and cancer of the colon and rectum is the fourth most common cancer in the world and the second most common cancer for both males and females. About 1.93 million new cases and 1 million die of colorectal cancer. CNN is being used to identify and classify colon cancer using whole-slide imaging (WSI). Accurate identification of is diseases the prerequisite of the treatment. Our objective is to promote a system for detecting and classifying clone adenocarcinomas by applying the transfer learning technique of convolution neural networks. This paper proposed some transfer learning models for the automated detection and classification of colon cancer. Leveraging a dataset of 10,000 images categorized into 2 types of colon adenocarcinoma and benign colonic tissue. The study evaluates the performance of various Convolutional Neural Networks (CNNs), including VGG19, Resnet152V2, DenseNet201 and InceptionResNetV2. Each model is trained using transfer learning techniques with fine-tuning, optimizing for accuracy and loss. The results demonstrate the effectiveness of deep learning in enhancing diagnostic accuracy, with VGG19, Resnet152V2, DenseNet201, and InceptionResNetV2 among them DenseNet201 reaching 100% of accuracy.**

Keywords—Convolutional Neural Networks, colorectal cancer, transfer learning, fine-tuning, image classification, automated detection.

I. INTRODUCTION

Cancer is one of the leading causes of death worldwide, and the World Health Organization has ranked it as the second most important cause of death, following coronary heart disease. Among the many types of cancer, colon cancer has the highest incidence and death rate. Especially in affluent societies with over 940,000 new cases and almost 500,000 deaths annually, this disparity underscores the urgent need for efficient diagnostic and therapeutic options. This is particularly true for underdeveloped regions, where sources of cancer detection are limited.

Cancers occur via the uncontrolled proliferation of cells, caused by genetic mutation leading to malignant tumors. Colon cancer frequently starts as non-diagnostic polyps, from which adenocarcinomas, the most frequent type of colon cancer, can arise. Since risk factors, including age, genetics, and lifestyle, influence its development, routine screening is of crucial importance, especially among those with hereditary risk factors.

Despite advancements in remedy, many economically developed countries are inadequately geared up to provide necessary pathology offerings, with the best 26% of such international locations having essential diagnostic tools available. This underscores an urgent want for investments in healthcare infrastructure, such as laboratory abilities and low-cost diagnostics, to mitigate cancer's impact on susceptible populations. Currently detecting the cancer is extremely time-consuming and labor-expensive. And there is always a lack of skilled pathologists. Recent trends in computer science, in particular gadget getting to know machine learning (ML) and deep learning (DL) getting to know, offer promising options for boosting most cancer detection and analysis. These technologies leverage computational strength to investigate histopathological photos greater efficaciously than traditional strategies, aiming to offer faster detection with more accuracy.

Our study's primary objective is to suggest a Deep Convolution Neural Network with transfer learning to investigate colon cancer by analyzing digitized pathology images.

II. METHODOLOGY

A. Dataset Description

The LC25000 dataset contains 25,000 color images with 5 classes of 5,000 images each. All images are 768 x 768 pixels in size and are in JPEG file format. The 5 classes are: colon adenocarcinomas, benign colonic tissues, lung adenocarcinomas, lung squamous cell carcinomas, and benign lung tissues. We took only two category images of colon adenocarcinomas, and benign colonic tissues for classification of colon cancer as rest three are lung-related diseases and we are working with colon cancer. Each category contains 5,000 images. Figure 1 illustrates some sample of LC25000 dataset images of colon adenocarcinomas, and benign colonic tissues. The dataset of LC25000 is available on this link: https://www.kaggle.com/datasets/antrnduy/lc25000.

B. *Preprocessing:*

In this study, we applied a combination of data pre-processing and augmentation techniques to improve model robustness and generalization. The images were preprocessed by normalizing pixel values from the range [0, 255] to [0, 1], ensuring efficient and stable training. For data augmentation, we used random transformations including zoom (±20%), shearing (±20%), rotation (±30^0), horizontal and vertical shifts (±20%), and horizontal flipping to artificially expand the dataset and introduce variability, helping the model become invariant to scale, orientation, and positioning changes. A 20% validation split was incorporated in both the training and testing data generators to prevent over-fitting and ensure reliable performance evaluation. These techniques, implemented via Keras's **ImageDataGenerator**, significantly enriched the training data, allowing the model to better generalize to new, unseen data and perform robustly in real-world scenarios.

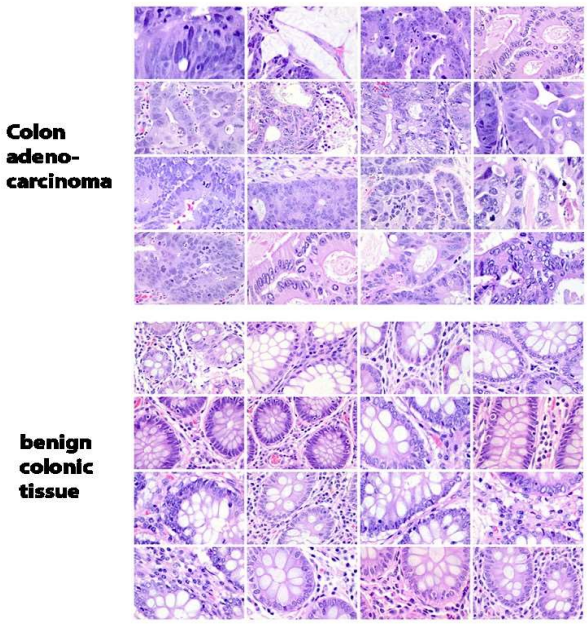

Fig 1: LC25000 dataset samples of two different categories of clone cancer.

C. *Model Architectures:*

VGG19: VGG19, proposed by Simonyan and Zisserman in 2014, is a deep convolutional neural network designed for large-scale image classification tasks. The architecture consists of 19 layers, including 16 convolutional layers and 3 fully connected layers, with a simple and uniform structure of 3x3 convolution filters stacked on top of each other. VGG19 was trained on the ImageNet dataset, which consists of over 1 million images across 1,000 object categories. The model achieved a top-5 accuracy of 92.7% on the ImageNet validation set, demonstrating its strong performance in image classification. Due to its simplicity and depth, VGG19 has been widely used as a baseline for various transfer learning tasks. The model's uniformity allows it to be easily extended for different architectures or fine-tuned for new tasks.

ResNet152V2: ResNet152V2, introduced by He et al. in 2016, is a deep residual network designed to address the problem of vanishing gradients in very deep networks. The ResNet architecture introduced the concept of residual connections, which allow gradients to flow more easily through the network. ResNet152V2 is an improved version of the original ResNet, featuring 152 layers with optimized batch normalization and activation functions. Trained on the ImageNet dataset, ResNet152V2 achieved a top-5 accuracy of 96.43% on the ImageNet validation set, making it one of the highest-performing models for image classification. The model's residual connections make it particularly effective for deep networks, allowing it to be trained efficiently without overfitting. The model has 152 layers and has become a standard for tasks requiring very deep architectures due to its strong performance and efficient training.

InceptionResNetV2: InceptionResNetV2, proposed by Szegedy et al. in 2016, combines the strengths of both the Inception architecture and ResNet's residual connections. The Inception module uses multiple parallel convolutions with different kernel sizes to extract features at various scales, while the residual connections from ResNet allow for more efficient gradient flow through the network. InceptionResNetV2 consists of 164 layers and was trained on the ImageNet dataset. It achieved a top-5 accuracy of 96.6% on the ImageNet validation set, offering an excellent trade-off between model size, depth, and performance. The model is well-suited for large-scale image classification tasks and has been widely adopted for transfer learning and fine-tuning applications in various domains. Its efficient architecture allows for faster training while maintaining high performance, especially in complex image recognition tasks.

DenseNet201: DenseNet201, introduced by Huang et al. in 2017, is a deep convolutional network that introduces dense connections between layers, where each layer receives input from all previous layers. This dense connectivity pattern helps alleviate the vanishing gradient problem and improves feature reuse. DenseNet201 has 201 layers and was trained on the ImageNet dataset, achieving a top-5 accuracy of 96.4% on the ImageNet validation set. DenseNet201 has fewer parameters compared to traditional architectures like ResNet and VGG due to its dense connectivity, making it more efficient while still maintaining high performance. The model's key feature is the dense block, where each layer concatenates its output with the outputs of all previous layers, promoting feature reuse and improving the model's representational power. DenseNet201 has proven to be highly effective in both image classification and transfer learning tasks.

D. Training Process

In this study, we utilized the pre-trained models **VGG19**, **ResNet152V2**, **InceptionResNetV2,** and **DenseNet201**,

excluding their top fully connected layers, to adapt them for our specific dataset. The convolutional base of each model served as a feature extractor. Fine-tuning was performed by unfreezing selected layers, allowing the deeper layers to adapt to the new task while preserving the low-level features learned from the ImageNet dataset. Specifically, the following layers were unfrozen for fine-tuning:

- *block5_conv1* for **VGG19**,
- *block8_4_conv* for **InceptionResNetV2**,
- *conv5_block1_preact_bn* for **ResNet152V2**
- *conv5_block1_0_bn* for **DenseNet201**.

To extend each model for multi-class classification, we appended the following layers:

- A flattened layer,
- A dense hidden layer with 256 units and ReLU activation
- A final output layer with sigmoid activation.

The models were trained using the RMSprop optimizer with a low learning rate of $1x10^{-5}$ to ensure minimal disruption to the pre-trained weights. A custom learning rate scheduler was employed to keep the learning rate constant for the first five epochs, followed by exponential decay. Early stopping was implemented to terminate training if the validation loss did not improve for four consecutive epochs. Additionally, model check-pointing was utilized to save the best model based on validation accuracy.

These strategies, including selective fine-tuning, a carefully designed architecture, and efficient training techniques, ensured that the models generalized effectively to the new task while retaining valuable knowledge from the ImageNet pre-training.

III. Results and Evaluation

A. Evaluation Matrix

To evaluate the performance of our deep learning models, we employed the precision, recall, and F1 score metrics, which are standard measures in classification tasks. These metrics are derived from the confusion matrix of each model, specifically focusing on the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). By using confusion matrix values we can generate precision, recall, and F1 scores to understand how good the model is, and get ideas about type 1 err and type 2 err. This illustrates how good the model is for predicting the correct value, and how bad to detect the correct value.
Fig 2 illustrates all model's confusion matrix and Table 1 conveys all model's evaluation metrics.

Table 1: Model evaluation metrics

Model	Precision	Recall	F1 Score
DenseNet201	100	100	1.00
InceptionResNetV2	100	99.9	99.9
ResNet152V2	99.9	100	99.9
VGG19	99.8	100	99.9

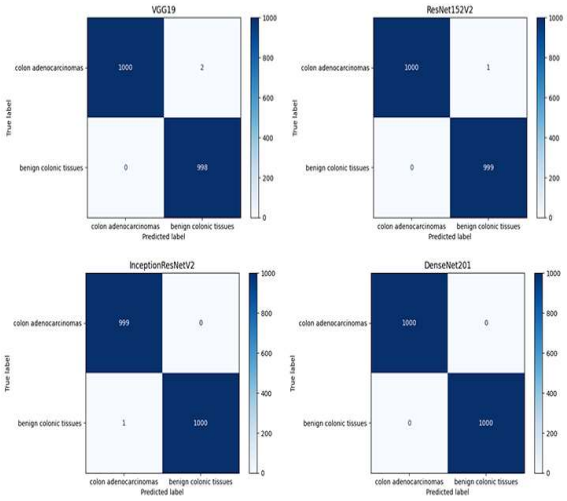

Fig 2: Models confusion matrix

B. Performance Matrix

The performance of the fine-tuned pre-trained models was evaluated for gastrointestinal disease classification. Table 1 presents the key metrics, including training accuracy, training loss, validation accuracy, and validation loss. DenseNet201 achieved the highest validation accuracy of 100% with a low validation loss of 0.0004, while InceptionResNetV2 and ResNet152V2 both achieved validation accuracies of 99.95% with minimal losses. VGG19 performed slightly lower with a validation accuracy of 99.90%.

As the models reaches too much accuracy in validation test, so we then use cross validation test and in cross validations we also got about the same accuracy as the validation accuracy. These results demonstrate the robustness and effectiveness of the models in accurately classifying gastrointestinal diseases, underscoring the potential of deep learning in medical diagnostics. Table 2 presents the model's performance metrics (Training Accuracy, Training Loss, Validation Accuracy, Validation Loss) Figure 2 and Figure 3 illustrate the accuracy curve of 7 epochs and the loss curve of 7 epochs respectively.

Table 2: Model performance metrics

Model	Training Accuracy %	Training Loss	Validation Accuracy %	Validation Loss
DenseNet 201	99.71	0.0085	100	0.0004
Inception ResNetV2	99.60	0.0127	99.95	0.0033
ResNet15 2V2	99.27	0.0215	99.95	0.0012
VGG19	99.76	0.9976	99.90	0.0017

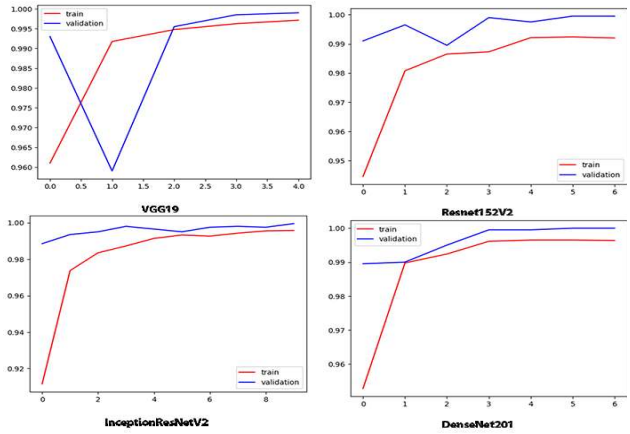

Fig 3: Models accuracy graph

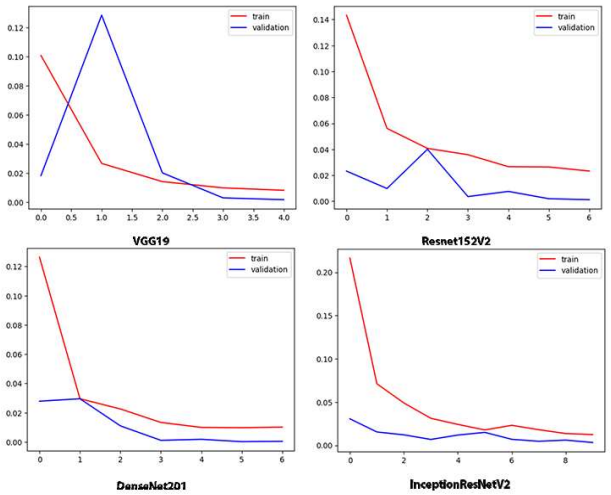

Fig 4: Models loss graph

IV. Discussion

This study demonstrates the effectiveness of transfer learning with deep CNNs for colon cancer detection. Pre-trained models like VGG19, ResNet152V2, DenseNet201, and InceptionResNetV2 were fine-tuned for classification tasks. DenseNet201 showed the best performance, likely due to its dense connectivity.

By DenseNet201 we got precision, recall, and f1 all 100%. Though all models perform well. However, the study is limited by the binary classification of colon adenocarcinoma and benign tissues, and the model's interpret-ability remains a challenge. Future work should focus on multi-class classification, model explainability, and dataset expansion for better generalization and clinical implementation.

V. Conclusion

This research highlights the potential of deep CNNs for automated colon cancer detection. DenseNet201 outperformed other models, providing high accuracy in identifying colon adenocarcinoma. While the results are promising, future research should address dataset limitations and enhance model interpretability. The integration of AI in clinical diagnostics could significantly improve early detection and treatment planning for colon cancer.

VI. References

[1] A. K. Gupta and S. Kumar, "Recent advances in the molecular pathogenesis of colon cancer," *Journal of Cancer Research and Clinical Oncology*, vol. 142, no. 6, pp. 1267-1276, 2016.

[2] F. Chen, S. Zhang, and Y. Wu, "The role of dietary fiber in colon cancer prevention," *Nutrition and Cancer*, vol. 70, no. 5, pp. 704-712, 2018.

[3] S. R. Sharma, A. Kumar, and V. K. Gupta, "Molecular biomarkers in early diagnosis of colorectal cancer," *BioMed Research International*, vol. 2018, Article ID 9730278, 2018.

[4] L. Smith, "Screening for colon cancer: Evidence-based guidelines," *American Journal of Preventive Medicine*, vol. 45, no. 2, pp. 95-102, 2020.

[5] G. D. James, P. K. Wilson, and R. B. Lee, "Colorectal cancer: Screening and treatment strategies," *Journal of Clinical Oncology*, vol. 35, no. 12, pp. 1562-1571, 2019.

[6] M. T. Tan, A. P. Goh, and K. N. Wong, "Colorectal cancer risk factors and prevention strategies in Southeast Asia," *Asian Pacific Journal of Cancer Prevention*, vol. 21, no. 4, pp. 931-937, 2020.

[7] K. Y. Oh, J. H. Lee, and M. J. Park, "Impact of genetic mutations in colon cancer prognosis," *Clinical and Experimental Medicine*, vol. 20, no. 3, pp. 357-365, 2019.

[8] D. H. Sweeney, R. M. McLoughlin, and L. W. Rhoads, "Immunotherapy in colon cancer: A new horizon," *Nature Reviews Clinical Oncology*, vol. 17, pp. 153-167, 2020.

[9] M. C. Yang, "Chemotherapy resistance in colon cancer," *Current Drug Resistance*, vol. 14, no. 5, pp. 537-544, 2018.

[10] B. F. Zhang, T. X. Chen, and P. Wang, "Role of gut microbiota in colon cancer and its therapeutic implications," *Frontiers in Microbiology*, vol. 11, p. 385, 2020.

[11] Y. S. Lee, M. H. Kwon, and J. S. Ahn, "Colon cancer metastasis and its microenvironment," *Journal of Cancer Metastasis and Treatment*, vol. 5, no. 1, pp. 12-19, 2019.

[12] T. I. Hashimoto, Y. Kato, and T. Horiguchi, "Colon cancer risk and lifestyle modifications," *Cancer Science*, vol. 108, no. 8, pp. 1572-1581, 2019.

[13] E. A. Anderson, "The role of colonoscopy in colon cancer prevention," *Journal of the American Medical Association*, vol. 322, no. 10, pp. 1012-1020, 2020.

[14] N. S. Patel and M. K. Gupta, "Molecular-targeted therapies in colorectal cancer," *OncoTargets and Therapy*, vol. 13, pp. 945-955, 2020.

[15] L. L. Tan, "Endoscopic management of colorectal cancer: New approaches," *Gastrointestinal Endoscopy Clinics of North America*, vol. 30, no. 4, pp. 609-622, 2020.