Migration Letters

Volume: 21, No: S11 (2024), pp. 449-463

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

Classification Of Colon Cancer Using Deep Learning Techniques On Histopathological Images

Ammar Ahmad Khan[1], Muhammad Arslan[2], Ashir Tanzil[3], Rizwan Abid Bhatty[4], Muhammad Asad Ullah Khalid[5], Ali Haider Khan[6]

Abstract:
Colon cancer is widely spread and deadly type of cancer in humans. Smoking contributes to the growth of lung cancer, which leads to an unbalanced diet, which can lead to colon cancer. Identification of this type of cancer depends largely on the histological diagnosis. One of the most important research goals is colon cancer identification and classification, especially in the field of health systems. The proposed model suggests a deep learning model based on histopathological images for the identification of colon cancer. The proposed model introduces a novel approach utilizing Vision Transformer (ViT) and a new ViT version known as Swin Transformer. The model achieves high accuracy as benign or malignant by using a modified version of the Swin Transformer model. A comparative analysis of several models Swin Transformer, ViT, ResNet-101 and a modified Swin transformer is also presented in this paper. LC25000 dataset is used for training phase and testing of the proposed model. The final results demonstrate that the test model's accuracy is 99.80% using the modified version fo Swin Transformer model, 99.64% using the Swin Transformer model, 99.36% using the ViT model, and 98.27% using ResNet-101.

Keywords: *Colorectal cancer, histopathological images, deep learning, Swin Transformer, cancer detection, image classification.*

1. Introduction

According to the World Health Organization, cancer is the biggest cause of death globally, with an estimated 28.4 million cases anticipated to be diagnosed worldwide by 2040 [1]. Abnormal weight gain, inactivity, heavy alcohol use, and poor diet are additional risk factors for cancer [2]. Age is a major non-preventable risk factor; in the United States, those 50 years of age or older account fo[1]r 87% of cancer cases [3]. Cancer is not a single disease, but rather a broad category of illnesses. Colon cancer is one of the deadliest cancers and is responsible for around 10% of cancer-related deaths worldwide [4]. Minimally invasive techniques, like histopathology, are necessary for precise disease detection and better treatment quality, since accurate classification of these tumors could not always be achievable utilizing non-invasive methods [5]. Initially, the patient either exhibits no symptoms or very little and when the patient experiences symptoms, the cancer has already progressed, and it is too late to treat the condition. Early detection of colon cancer is challenging to achieve without thorough testing, including tissue inspection, MRI, CT scans, and ultrasounds.

In recent years, machine learning (ML) has provided crucial support in medical image processing [6]. Machine learning technology is used to detect, diagnose and classify various diseases based on medical images. This technology helps develop automated disease diagnosis models using patient images. Convolutional Neural Networks (CNNs) are

[1]Department of Computer Science, NAMAL University, Mianwali,42250, Punjab, Pakistan.
[2]Faculty of Computer Science, Lahore Garrison University, Lahore, 54000, Pakistan.
(Corresponding Author)
[3]Department of Computer Science, Abasyn University Islamabad Campus, Islamabad,44000, Pakistan.
[4]Department of Business, University of The Cumberlands , Williamsburg, KY 40769, USA.
[5]Faculty of Computer Science, Lahore Garrison University, Lahore, 54000, Pakistan.
[6]Faculty of Computer Science, Lahore Garrison University, Lahore, 54000, Pakistan.

artificial neural networks with deep learning capabilities that are inspired by the structure of the visual brain of animals and are used to evaluate visual images [7]. Because of their hierarchical nature as multilayer perceptrons, they can construct complex patterns from smaller ones.

CNNs require less pre-processing than standard image classification techniques because the network optimizes filters on its own [8]. As a result, less substantial human intervention and hand-engineering are required. Because CNNs are fully connected, they can be vulnerable to overfitting; however, regularization strategies like weight decay and dropout help reduce this risk. CNNs are able to efficiently cover the whole visual field because of the overlapping receptive fields of individual neurons [9].

Histopathological images, which are crucial to the diagnosis of diseases like cancer, are microscopic images of stained tissue samples used to investigate disorders at the cellular level and a Vision Transformer (ViT) uses a sliding window controlled by stride and patch size to split an image into non-overlapping or user-defined overlapping patches [10]. These patches become flattened vectors after being vectorized. After that, a linear function known as a dense layer is applied, producing the vectors Z1, Z2,..., Zn. To improve accuracy, positional encoding is applied to these vectors in order to capture both position and content. In this research work, a thorough analysis of various models including Resnet101 [11], Vit , Swin transformers and modified version of Swin transformer is presented to evaluate the efficiency in detecting colon cancer. The purpose of this research work is to create a model that uses deep learning techniques to extract features, classify, and forecast histopathological images, assisting medical professionals and researchers in the identification of colorectal cancer. For early detection and remedy strategies, image processing techniques are normally used, and the identification of accepted and environmental factors may be very crucial for the prediction of colorectal cancer. Therefore, the technique of early prediction and detection of colorectal cancer should propose which is simple, cost effective and time saving, it will result in positive consequences for colorectal most cancers prognosis and prediction. Early detection and prediction of colorectal cancer can even play an essential role inside the prognosis manner and improve the patient's survival rate too. We used LC25000 dataset file. This system proposed to detect the colorectal cancer detection. First of all, input Histopathological image of colon. Then we preprocessed it, in preprocessing filtering is required. The primary purpose of filtering is smoothing of an image. We used a Gaussian filter for our histopathological images because it work efficiently for our datasets. Then we used image classification using Swin Transformer model and did comparative study with other deep learning techniques. We test and train the histopathological images and detect cancer from colon. After implementation we get 99.80% accuracy using modified last layer Swin transformer

The research paper's remaining sections are separated as follows: Section 2 presented the overview of previous work done on this domain. Section 3 discusses methodology including all the image processing techniques, deep learning models. In section 4, results are displayed. Section 5 covers the research's conclusion.

2. Literature Review:

Colorectal cancer the growth of a tumor that starts from the inside wall of the colon or rectal cavity . It is the most dangerous disease worldwide. Early identity of colorectal cancer is a difficult problem due to most cancer cellular composition, in which several of the cells overlap each other. This section presents the previous work done on colon cancer detection, In [2], the combination of three approaches were proposed for the early diagnosis of LC25000 dataset histology images, each having two systems. An improvement has been made to histological images, with damaged areas showing more contrast. The PCA method was used to eliminate redundant and unneeded features in order to reduce high dimensionality while retaining key characteristics, as the GoogLeNet and VGG-19 model of each system provided large-scale characteristics. The first method for using ANN to diagnose the histology images in the LC25000 dataset makes use of specific features from the VGG-19 and GoogLeNet models. The second approach makes use of ANN and combines VGG-19 and GoogLeNet features. While one system merge complex features

and subsequently decreased high dimensions, another system decreased size and merged. The third approach combines handcrafted features and ANN with fusion characteristics of CNN models VGG-19 and GoogleNet.

Moreover in [1], the authors presented a model using LC25000 histopathology images to automatically identify lung cancer subtypes, including colon adenocarcinoma and colon benign. EffcientNetV2, a large, micro model deep learning based architecture, is based on the ideas of progressive learning and combined scaling. The EffcientNetV2-L model outperformed the previous techniques The authors employed gradCAM to generate visual saliency maps that allow to effectively identify the critical areas in the test set histopathological pictures where the models focus greater attention when predicting cancer subtypes. The pathologists who create these visual saliency maps might find it easier to develop more effective treatment plans. As a result, the suggested pipeline can be used in hospitals for completely technologically advanced, explainable colon cancer identification from histopathology images.

A new computationally feasible model was presented for the rapid and precise identification of malignancies in colon regions as an alternative to the existing methods of cancer detection [12]. The LC25000 dataset is utilized in the training, validation, and testing stages of this project. The cyclic learning rate employed in these approaches improves their accuracy while maintaining their processing efficiency. This is simple and efficient, which helps the model converge more quickly. Furthermore, several pre-trained models for transfer learning are used and compared with the CNN which is being proposed from the beginning. It is found that the proposed model provides improved accuracy, reducing the impact of inter-class variations between lung adenocarcinoma and lung squamous cell cancer. By putting the suggested strategy into practice, total accuracy was raised to 97% and computational efficiency was demonstrated in comparison to alternative methods.

A deep learning approach-based strategy for automatically classifying lung and colon cancer was presented in this work [13]. The histological image of benign cancer, malignant cancer, and normal tissue is the subject of the categorized image. To demonstrate classification on 25,000 histopathology images, VGG16, and Contrast Limited Adaptive Histogram Equalization (CLAHE) were used. According to the simulation findings, this model can classify data with an average accuracy of 98.96%. When utilizing CLAHE, the system performance is consistent across all epoch settings and exhibits a greater detection accuracy than when without using CLAHE. The comparison analysis reveals that the suggested approach performs better than some earlier research. It is envisaged that this suggested approach can assist medical professionals in autonomously identifying cancer, with minimal costs, high accuracy, and quick processing on big datasets.

In [14], the authors presented an ensemble classifier that makes use of three different methodologies: logistic regression (LR) model, a support vector machine (SVM), and random forest (RF). To produce the ensemble classifier, predictions of each one are combined using a majority decision method. Two distinct techniques have been used to extract complex features from colon images: local binary pattern (LBP) and VGG16. The initial appropriate parameters for the group classifier are obtained through the combination of the LBP and VGG16 features. The colon histopathology (LC25000) datasets are used to validate the proposed technique.

In [15] paper suggests a novel machine-learning method for identifying lung and colon malignancies. To improve diagnostic accuracy, a machine learning-based hybrid classification model is to be developed. Twenty-five thousand color histopathological imaging samples colon cell tissues, showing a diagnosis or absence of cancer (adenocarcinoma), make up the LC25000 dataset. The VGG-16 model that has been pre-trained is used to extract image features. Three classification using machine learning based techniques are used to identify the type of cancer: K-Nearest Neighbor (KNN), Random Forest (RF), and Stochastic Gradient Descent (SGD). CNN-SGD performed the best according to assessment metrics during the evaluation of model, which used a 10-fold cross-validation technique.

Table1 Comparison Table for Previous Models

Author	**Models/Methods Used**	**Limitation**
M. Al-Jabbar et al. [2] 2023	GoogLeNet, VGG-19, ANN, PCA	Need for large and diverse datasets to generalize findings, and computational complexity associated with high-dimensional feature processing.
S. Tummala et al. [1] 2023	EffcientNetV2 (large, medium, small), gradCAM	Extensive computational resources for training deep learning models is required, dependency on high-quality histopathology images,
M. A. M. Provath et al [12] 2023	CNN from scratch, transfer learning models, cyclic learning rate	Efficiency gains need to be validated on different hardware setups, and the model's performance should be tested across diverse populations and imaging conditions to ensure generalizability.
S. Hadiyoso et al [13] 2023	CNN (VGG16), CLAHE	The models' applicability and reliability in diverse clinical settings require extensive validation. The computational cost and processing time for large datasets also need to be considered.
O. Singh et al [14] 2023	RF, Support SVM, VGG16, Local Binary Pattern (LBP), ensemble classifier	combining multiple classifiers may lead to increased training and prediction times.
Chilyatun Nisa et al [15] 2023	VGG-16, SGD, RF, K-Nearest Neighbor (KNN)	The model's performance needs validation across different datasets and clinical settings.

3. Methodology

This section presents the complete overview of the models and ML techniques used in this research to identify the LC25000 dataset's histopathology images. The first step is pre-processing of the LC25000 training datasets. Filtering is applied on datasets to make dataset images smooth and remove noise [16]. The primary purpose of filtering is smoothing of an image or suppressing the high frequency in the image. There are many types of filters used for different purposes according to the specifications [16]. We used a Gaussian filter for our histopathological images because it will work efficiently for our datasets [17]. After

preprocessing, features are extracted using the Vision Transformer model. Image feature extraction is an important step that uses algorithms and techniques to identify and extract various preferred parts or shapes from the input image. The flow of the proposed model includes input image, image preprocessing, data partitioning, training and testing, feature extraction. and Classification

Figure 1: Flowchart of proposed model

Image Preprocessing

In the pre-processing, the input histopathology image is processed to improve the image quality, and at the same time, different operations are performed on the image. The details and specific data are improved. This improved version will facilitate the subsequent stages of any robotic system. Therefore, it is useful to perform some pre-processing operations. Data pre-processing is used to remove noise, if any and give helpful information from these images [18]. The steps of image pre-processing are: read the image, adjust the image size, remove noise, also known as Denoise, segmentation in which we are separating the background from foreground objects, and Morphology which is smoothing the edges [19]. After this, we portioned our data into two parts, one for Testing and training. Our model will check the testing data and give us some results, whether malignant or benign

Figure 2: The flow diagram of the proposed model

Data Partitioning

After data augmentation of images, we split our dataset utilizing a 75% ratio for Training and 25% for Testing and 20% splitting the date for validation. According to the standard rule for best training results, after training images, we pass the model to the testing phase to check the accuracy and accurate positive error results.

Training and Testing

Different parameters are set for best weights learning, when we train our model. Input image size, Learning Rates, Optimizer, Weight decay, Momentum, Epochs, Batch size these all parameters are set in following ways.

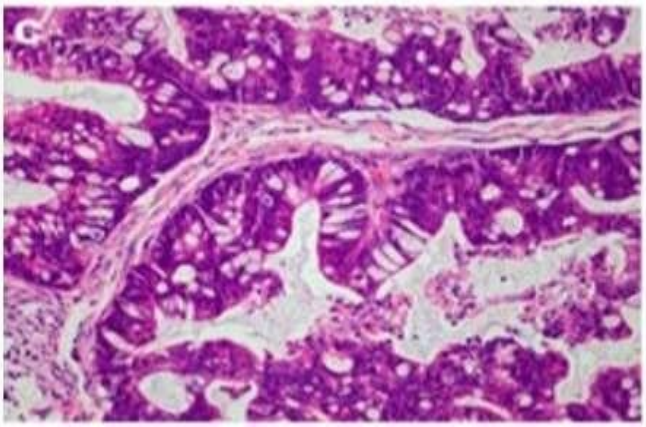

Figure 3: Input histopathological image of Colon

Table 2: Training Parameters for Model

1	Input Size	224,224,3
2	Learning Rates	0.01

3	Optimizer	Adam
4	Weight Decay	0.01
5	Momentum	0.9
6	Batch Size	32
7	Epochs	30(early stopping)

The above table shows the training parameters which are used in this system for validation purpose.

Input Size

In training the input size of Histopathological images of Colon Cancer is 224,224,3.

Optimizer

In optimization we used the Adam [20]. Adam is a customized adaptable learning rate optimization method created for deep neural network training. Adam is used to calculate learning rates of different parameters individually.

Learning Rates

When we train the model, we set learning rate LR 0.01 [21]. An optimization algorithm that determines the number of steps at every repetition while aiming for the minimum value of a loss function may use the learning rate as a tuning parameter. In setting a learning rate, there's a trade-off between the speed of convergence and overshooting.

Weight Delay

While training our model we set weight decay to 0.01. It is a deep learning regularization method. It functions by introducing a penalty period. to the value feature of a neural community, which causes the weights to shrink during backpropagation [22]. This lessens the possibility that the network will explode the gradient problem and overfit the training set.

Momentum

We use 0.9 as momentum while training [23]. The basic concept of momentum in machine learning is to increase the speed of training.

Batch Size

The batch size is 32. The quantity of samples processed prior to the model being updated is referred to as the batch size [24]. The size of a batch must be more than or equivalent to one and not exactly or equivalent to the number of samples in the trained dataset.

Epochs:

We are using 30 epochs for training our data and we used early stopping in each epoch, so it automatically stops. An epoch is the indication of the number of passes of the whole training datasets which is completed by using any machine algorithm.

Feature Extraction:

Image feature extraction is an important step that uses algorithms and techniques to identify and extract various preferred parts or shapes from the input image. In our system, CNN and Transformer done the feature extraction by themselves.

Classification

We classify our dataset into two groups. One is malignant and the other is benign so when our system shows 0 it means it is malignant and when shows 1 then it means it is benign.

General Model Vision Transformer

We give the image as input to Vision Transformer, then we split an image into patches, but these patches do not overlap automatically, but the user can do overlapping through a sliding pane that shifts multiple pixels every time, and the user can do this by applying stride. A minor stride results in a more significant count of the patches. To split an image user must specify two arguments. One is patch size, and the other is stride. Then flatten the patches by applying vectorization. To vectorize something is to transform the tensor into a vector, and tensor is how much order for an image like if we have a color image, then it is RGB means three channels, so it means tensor is of order 3The vectors are d1*d2*d3*1 if the patches are d1*d2*d3 tensors, and after applying vectorization, we get output x1, x2…, xn. After vectorization, we apply a dense layer; the dense layer is a linear function. Dense layer shared parameters. Parameters are W and b learned from training data. So, after applying the dense layer, the resultants are Z1= Wx1+b, Z2= Wx2+b…, Zn= Wxn+b depending upon the number of patches. Positional encoding is also used because when the input image is split into patches, we give each patch a position of an integer from 1 to n. In positional encoding, a vector is mapped to an integer. Given that the vector's shape is same to Z's, add positional encoding to Z1, Z2…, Zn vectors. Thus, a Z vector now stores a patch's position as well as its content. Without positional encoding, the accuracy decreased by 3%. Another layer is used, named CLS token for classification, and the output of that layer is vector Zo. Fit Zo to Zn vector to a multi-head self-attention layer. An extra learnable embedded is attached to predict the class of an input image. After applying a dense layer, which also produces n+1 vectors, the multi-head self-attention output is a sequence of n+1 vectors. We use multiple layers of the multi-head and dense layer according to the situation, and these layers combined are known as transformer

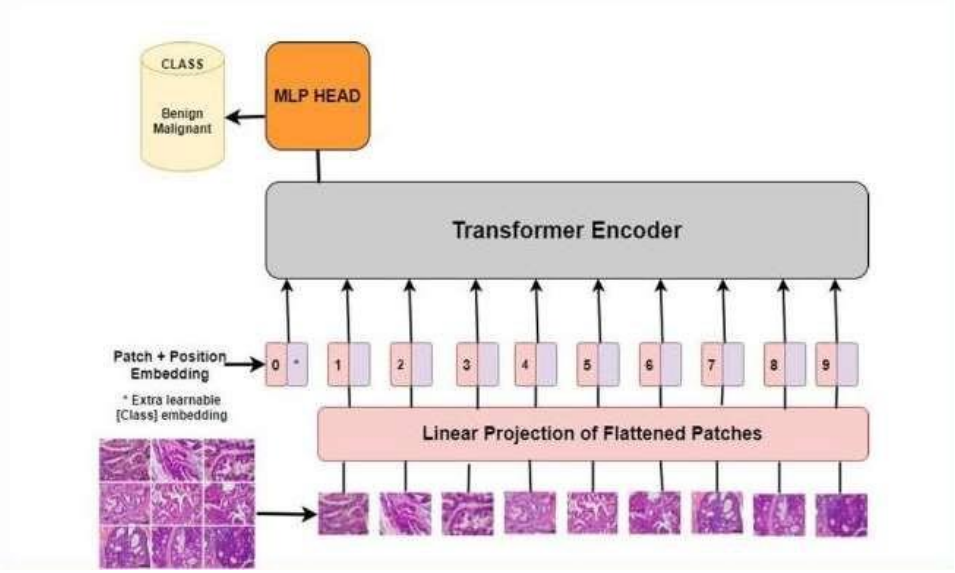

Figure 4: An illustration of Vision Transformer (ViT) architecture.

encoder networks. The output of the transformer encoder is Co…, Cn, but for classification, we do not need C1…, Cn we only need Co to classify. The feature vector that is obtained from an image is Co. Feed Co to a SoftMax classifier and SoftMax classifier outputs vector p. P is the number of dataset classes. The architecture of the Vision Transformer (ViT) is illustrated in Figure 4.

Swin Transformer

The swin model employs a shifted window method, which overcomes the vision transformer's limitations of recording huge fluctuations in the scale of visual entities and high pixel resolution in an image [25].

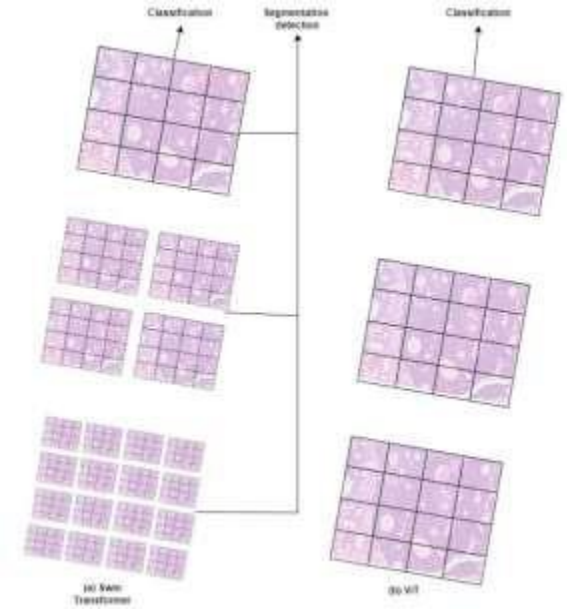

Figure 5: Hierarchical feature maps

The idea of Swin transformer is shown in the above figure. On the right side of above figure is vision transformer and we can see that the number of patches is kept the same throughout different layers. Hence, when the model moves further into the transformer layers, the Swin transformer increasingly combines neighboring patches instead of starting with an equal number of small patches with many windows in each layer. The complexity of the procedure is linear. This is different from previous ViT approaches with quadratic computational complexity, since self-attention needs to be applied globally to all patches.

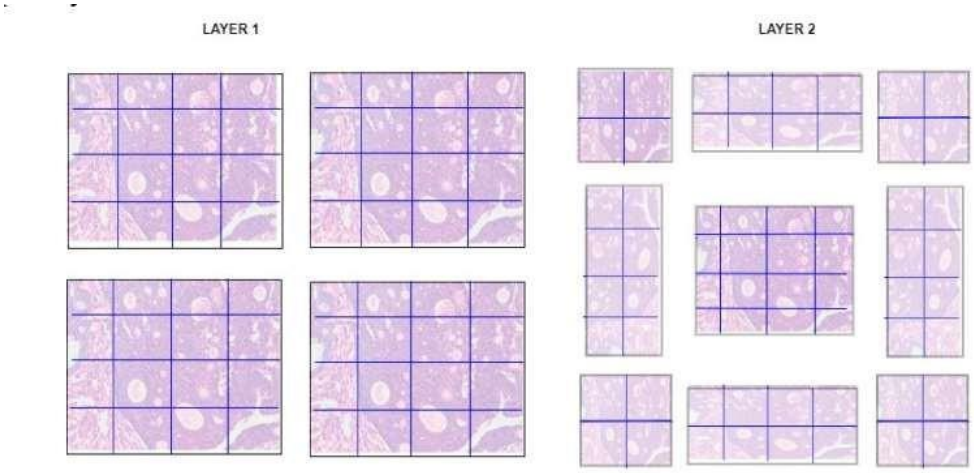

Figure 6: Core idea of shifting window

The essential concept of shifted windows is depicted in the diagram above. The blue area is a patch, and the black outline indicates the window that groups neighboring patches together. Each window in the first layer is 4*4, resulting in 16 patches per window. After performing self-attention locally within each window, the following layer shifts the windows to the bottom right, resulting in a new set of windows in the right. So L2 is the new resultant of a new group. Since self-attention is calculated locally within the non-overlapping windows that divide an image, the computational complexity is achieved in a linear fashion. meaning self-attention is performed locally in the black outlined region because the window size is fixed the computational complexity is linear with respect to the

image size. The swin transformer model's architecture is depicted in the accompanying figure. In the patch portioned stage, the input image is first divided into patches, and then each patch is projected using a linear layer in the first stage. There are now numerous stages, each of which is made up of a patch merging layer and multiple swin transformer blocks. As the network expands, the patch merging block is utilized to minimize the quantity of tokens. The first layer of patch merging concatenates the characteristics of every group of adjoining patches before applying a linear to them.

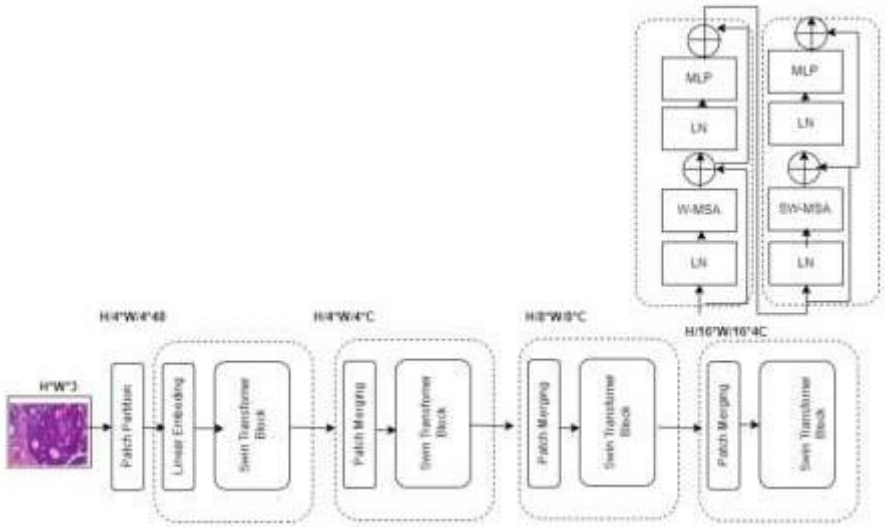

Figure 7: Architecture of swin Transformer

We can see on the right side of above figure, the building blocks of the swin transformer block in general it is similar to the block in transformers. The WSMA donates a regular self-attention module and the SWMA donates the attention layer with a shifted windows configuration. The shifting of the window is performed here in the SW-MSA layer. A development of the Vision Transformer (ViT) intended to overcome particular constraints in image processing applications is the Swin Transformer. Images in ViT are broken down into 16x16 pixel patches. These patches are then converted into patch vectors and handled in a manner akin to that of word vectors in text transformers. Large patches are sufficient for this strategy to operate effectively in applications like picture categorization. However, ViT has trouble with tasks like semantic segmentation that call for precise pixel-level information. Using ViT to handle high-resolution photos produces unmanageably long sequences, which makes in-depth analysis difficult. To get around this, the Swin Transformer uses a hierarchical structure with shifted windows, which makes it more effective at processing images of various sizes.

Changes to the Swin Transformer's Final Layer:

To improve model accuracy, we made the following changes to the Swin Transformer's last layer:

1. A dropout layer was added, with a probability of 0.3.
2. 256 outputs from a linear layer that was appended.
3. Another Dropout layer with a probability of 0.5 was added.
4. Added a last layer for binary classification, producing a result of 0 for healthy cells and 1 for malignant cells.

ResNet-101

A deep convolutional neural network with 101 layers, ResNet-101 [11], tackles issues with

deep network training such disappearing gradients. ResNet makes use of skip connections and residual blocks to increase model accuracy and make training very deep networks easier. These characteristics improve the model's performance on training and testing data while also assisting in mitigating the degradation issue. Residual blocks are used in ResNet's deep residual learning to improve model accuracy. Even in very deep networks, higher performance is ensured by the network's ability to learn identification functions through skip connections. By merging outputs ferrorrlier layers with outputs from stacked layers, this lowers errors and permits efficient training of deeper networks.

4. RESULTS AND DISCUSSION:

Dataset

Dataset for our research work is LC25000 [26] which is abbreviated as Lung Colon and 25000 means 25000 color images are present. This dataset is divided into five classes having 5000 images in each class. This dataset is a combination of colon and lung cancer images, but we only need colon cancer images, so if we work on this, we will only take two classes: benign colonic tissue and colon adenocarcinoma. These are histological images; histology is the microscopic examination of a surgical or biopsy sample that has been treated and preserved on slides in order to research disease symptoms. The sections are stained with one or more colors to help view the different components of the tissue under the microscope.

Performance Evaluation

We are using the accuracy, sensitivity and specificity as the performance evaluation parameters.

Accuracy

The percentage of true positives or true negatives in the population is known as accuracy. the percentage of the total number of correctly anticipated outcomes. It can be calculated as [27]:

$$\mathrm{TN} + \frac{\mathrm{TP}}{\mathrm{Total}} \qquad (1)$$

Recall

The percentage of true positives that are accurately identified is called recall. It demonstrates how well the test can identify the benign or malignant class. It can be calculated as [28]:

$$\frac{\mathrm{TP}}{\mathrm{TP+FN}} \qquad (2)$$

Precision

The ratio of accurately anticipated positive observations to all predicted positive observations is known as precision. It can be calculated as [29]:

$$\frac{\mathrm{TP}}{\mathrm{TP+FP}} \qquad (3)$$

F1Score

It is the mean of precision and recall.. It can be calculated as [30]:

$$\text{f1 Score} = 2 * \frac{(\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})} \qquad (4)$$

Table 3: Performance evaluation parameter of model

Normal_ True	Normal_ False	Cancer_ False	Cancer_ True	Total	Accuracy	Precision	Recall	F1 Score

1236.0	15.0	12.0	1241.0	2504.0	0.989217	0.990423	0.988057	0.989239
1246.0	4.0	5.0	1249.0	2504.0	0.996406	0.996013	0.996808	0.99641
1248.0	2.0	3.0	1251.0	2504.0	0.998003	0.997608	0.998404	0.998006
1242.0	8.0	8.0	1246.0	2504.0	0.99361	0.99362	0.99362	0.99362

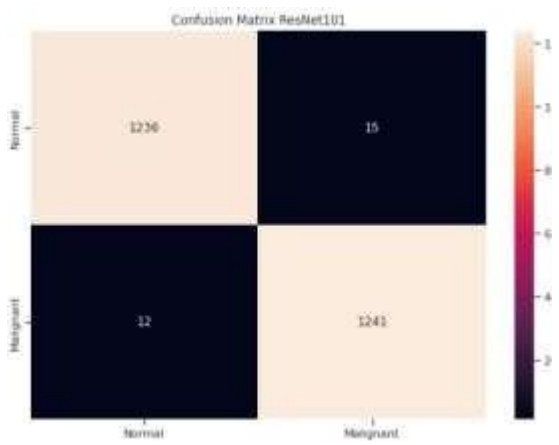

Figure 8: Confusion Matrix using ResNet 101

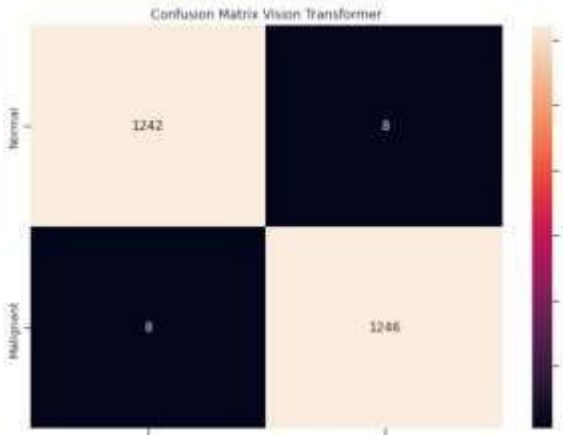

Figure 9: Confusion Matrix using Vision Transformer

Figure 10: Confusion Matrix using Swin Transformer for using Swin Transformer

Figure 11: Confusion Matrix

Modification Last Layer

Table 4: Comparison Table of Results

Comparison Result	ResNet-101	Vision Transformer	Swin Transformer	Swin Transformer Modified Last Layer
Accuracy	98.27%	99.36%	99.64%	99.80%
Precision	99.04%	99.36%	99.60%	99.76%
Recall	98.80%	99.36%	99.68%	99.84%
F1-Score	98.92%	99.36%	99.64%	99.80%
Normal_True	1236	1242	1246	1248
Normal_False	15	8	4	2
Cancer_False	12	8	5	3
Cancer_True	1241	1246	1249	1251

In case of Colorectal Cancer Classification using histopathological images, we are getting accuracy of 98.27% using ResNet_101, 99.36% using Vision Transformer, 99.64% using Swin Transformer, and 99.80% using Modified last layer Swin Transformer. The table shows the precision ratio of 99.04% for ResNet-101, 99.36% for Vision Transformer, 99.60% for Swin transformer and 99.76% for Modifies version of Swin transformer. The results show that ResNet-101 produces recall ratio of 98.80% ,99.36% recall for vision transformer, 99.68% for Swin transformer, 99.84% recall for modified version of Swin transformer.

Conclusion:

Colorectal cancer is one of the widespread and severe disease that creates a major public health problem, and it has a high death rate. In this regard, proper classification of colon cancer from, histopathological images are used towards achieving completely computerized diagnosis system for colon cancer detection. With this improvement of model and availability of data, the histopathologist can save their valuable time by using computer tools for colon cancer classification. In this thesis work, our main goal is to detect malignant and benign colon cancer classification from histopathological images by using Transformers model. The LC25000 dataset served as our training and testing set of 25000 histopathological images. In this proposed system we develop a modify layer of transformer using Swin transformer, using transformer architecture to detect colon cancer. This model is used to point out the colon cancer of patients.

References:

[1] S. Tummala, S. Kadry, A. Nadeem, H. T. Rauf, and N. Gul, “An Explainable Classification Method Based on Complex Scaling in Histopathology Images for Lung and Colon Cancer,” Diagnostics 2023, Vol. 13, Page 1594, vol. 13, no. 9, p. 1594, Apr. 2023, doi: 10.3390/DIAGNOSTICS13091594.

[2] M. Al-Jabbar, M. Alshahrani, E. M. Senan, and I. A. Ahmed, “Histopathological Analysis for Detecting Lung and Colon Cancer Malignancies Using Hybrid Systems with Fused Features,” Bioengineering, vol. 10, no. 3, p. 383, Mar. 2023, doi: 10.3390/bioengineering10030383.

[3] M. Sun, J. Wang, Q. Gong, and W. Huang, “Enhancing gland segmentation in colon histology images using an instance-aware diffusion model,” Comput. Biol. Med., vol. 166, p. 107527, Nov. 2023, doi: 10.1016/j.compbiomed.2023.107527.

[4] M. Murugesan, R. Madonna Arieth, S. Balraj, and R. Nirmala, “Colon cancer stage detection in colonoscopy images using YOLOv3 MSF deep learning architecture,” Biomed. Signal Process. Control, vol. 80, p. 104283, Feb. 2023, doi: 10.1016/j.bspc.2022.104283.

[5] E. A.-R. Hamed, M. A.-M. Salem, N. L. Badr, and M. F. Tolba, “An Efficient Combination of Convolutional Neural Network and LightGBM Algorithm for Lung Cancer Histopathology Classification,” Diagnostics, vol. 13, no. 15, p. 2469, Jul. 2023, doi: 10.3390/diagnostics13152469.

[6] C. Jones, D. C. Castro, F. De Sousa Ribeiro, O. Oktay, M. McCradden, and B. Glocker, "A causal perspective on dataset bias in machine learning for medical imaging," Nat. Mach. Intell. 2024 62, vol. 6, no. 2, pp. 138–146, Feb. 2024, doi: 10.1038/s42256-024-00797-8.

[7] W. El-Shafai et al., "Efficient classification of different medical image multimodalities based on simple CNN architecture and augmentation algorithms," J. Opt., vol. 53, no. 2, pp. 775–787, Apr. 2024, doi: 10.1007/S12596-022-01089-3/TABLES/4.

[8] M. Maree, T. Zanoon, A. Dababat, and M. Awwad, "Constructing a hybrid activation and parameter-fusion based CNN medical image classifier," Int. J. Inf. Technol., vol. 16, no. 5, pp. 3265–3272, Jun. 2024, doi: 10.1007/S41870-024-01798-X/TABLES/4.

[9] A. R. M Abbas, M Arslan, R A Bhatty, F Yousaf, A A Khan, "Enhanced Skin DiseaseDiagnosis through Convolutional Neural Networks and Data Augmentation Techniques," J. Comput. Biomed. Informatics, vol. 07, no. 01, 2024.

[10] G. M. S. Himel, M. M. Islam, K. A. Al-Aff, S. I. Karim, and M. K. U. Sikder, "Skin Cancer Segmentation and Classification Using Vision Transformer for Automatic Analysis in Dermatoscopy-Based Noninvasive Digital System," Int. J. Biomed. Imaging, vol. 2024, pp. 1–18, Feb. 2024, doi: 10.1155/2024/3022192.

[11] R. R. Wahid, C. Nisa', R. P. Amaliyah, and E. Y. Puspaningrum, "Lung and colon cancer detection with convolutional neural networks on histopathological images," 2023, p. 020020. doi: 10.1063/5.0114327.

[12] M. A. M. Provath, K. Deb, and K. H. Jo, "Classification of Lung and Colon Cancer Using Deep Learning Method," Commun. Comput. Inf. Sci., vol. 1857 CCIS, pp. 56–70, 2023, doi: 10.1007/978-981-99-4914-4_5.

[13] S. Hadiyoso, S. Aulia, and I. D. Irawati, "Diagnosis of lung and colon cancer based on clinical pathology images using convolutional neural network and CLAHE framework," Int. J. Appl. Sci. Eng., vol. 20, no. 1, pp. 1–7, 2023, doi: 10.6703/IJASE.202303_20(1).006.

[14] O. Singh and K. K. Singh, "An approach to classify lung and colon cancer of histopathology images using deep feature extraction and an ensemble method," Int. J. Inf. Technol., vol. 15, no. 8, pp. 4149–4160, Dec. 2023, doi: 10.1007/S41870-023-01487-1/TABLES/5.

[15] "CLASSIFICATION OF LUNG AND COLON CANCER TISSUES USING HYBRID CONVOLUTIONAL NEURAL NETWORKS | Nisa' | JUTI: Jurnal Ilmiah Teknologi Informasi." http://juti.if.its.ac.id/index.php/juti/article/view/1225 .

[16] N. S. Awarayi, F. Twum, J. Ben Hayfron-Acquah, and K. Owusu-Agyemang, "A bilateral filtering-based image enhancement for Alzheimer disease classification using CNN," PLoS One, vol. 19, no. 4, p. e0302358, Apr. 2024, doi: 10.1371/journal.pone.0302358.

[17] M. Cheng and M. Liu, "Image convolution techniques integrated with YOLOv3 algorithm in motion object data filtering and detection," Sci. Reports 2024 141, vol. 14, no. 1, pp. 1–13, Apr. 2024, doi: 10.1038/s41598-024-57799-0.

[18] F. L. de la Rosa, L. Moreno-Salvador, J. L. Gómez-Sirvent, R. Morales, R. Sánchez-Reolid, and A. Fernández-Caballero, "Improved Surface Defect Classification from a Simple Convolutional Neural Network by Image Preprocessing and Data Augmentation," pp. 23–32, 2024, doi: 10.1007/978-3-031-61137-7_3.

[19] A. Alqushaibi et al., "Enhanced Colon Cancer Segmentation and Image Synthesis through Advanced Generative Adversarial Networks based-Sine Cosine Algorithm," IEEE Access, pp. 1–1, 2024, doi: 10.1109/ACCESS.2024.3402262.

[20] A. Bhattacharya, B. Saha, S. Chattopadhyay, and R. Sarkar, "Deep feature selection using adaptive β-Hill Climbing aided whale optimization algorithm for lung and colon cancer detection," Biomed. Signal Process. Control, vol. 83, p. 104692, May 2023, doi: 10.1016/J.BSPC.2023.104692.

[21] A. T. Azar et al., "Automated System for Colon Cancer Detection and Segmentation Based on Deep Learning Techniques," Int. J. Sociotechnology Knowl. Dev., vol. 15, no. 1, pp. 1–28, Jul. 2023, doi: 10.4018/IJSKD.326629.

[22] S. Sakowitz et al., "Association of time to resection with survival in patients with colon cancer," Surg. Endosc., vol. 38, no. 2, pp. 614–623, Feb. 2024, doi: 10.1007/s00464-023-10548-2.

[23] Z. Gottschalk and S. A. Cohen, "Use of Circulating Tumor DNA to Guide Decision-making in Adjuvant Colon Cancer," Curr. Oncol. Rep., Jun. 2024, doi: 10.1007/s11912-024-01565-y.

[24] Y. G. C. Kelana, S. Rizal, and S. Saidah, "Classification of Histopathological Images of Colon Cancer Using Convolutional Neural Network Method," in 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), IEEE, Feb. 2023, pp. 821–826. doi: 10.1109/ICCoSITE57641.2023.10127795.

[25] H. Wang and P. Li, “Shifted window-based Transformer with multimodal representation for the systematic staging of rectal cancer,” Serv. Oriented Comput. Appl., May 2024, doi: 10.1007/s11761-024-00400-3.

[26] A. Merabet, A. Saighi, M. A. Ferradji, and Z. Laboudi, “Enhancing Colon Cancer Prediction in Histopathology with Integrated Deep Learning Models: A Comparative Study on the LC25000 Dataset,” in 2024 6th International Conference on Pattern Analysis and Intelligent Systems (PAIS), IEEE, Apr. 2024, pp. 1–7. doi: 10.1109/PAIS62114.2024.10541219.

[27] E. Bostanci, E. Kocak, M. Unal, M. S. Guzel, K. Acici, and T. Asuroglu, “Machine Learning Analysis of RNA-seq Data for Diagnostic and Prognostic Prediction of Colon Cancer,” Sensors, vol. 23, no. 6, p. 3080, Mar. 2023, doi: 10.3390/s23063080.

[28] S. Schwindt, L. Meisinger, B. Negreiros, T. Schneider, and W. Nowak, “Transfer learning achieves high recall for object classification in fluvial environments with limited data,” Geomorphology, vol. 455, p. 109185, Jun. 2024, doi: 10.1016/j.geomorph.2024.109185.

[29] K. T. Ahmed, H. Afzal, M. R. Mufti, A. Mehmood, and G. S. Choi, “Deep Image Sensing and Retrieval Using Suppression, Scale Spacing and Division, Interpolation and Spatial Color Coordinates with Bag of Words for Large and Complex Datasets,” IEEE Access, vol. 8, pp. 90351–90379, 2020, doi: 10.1109/ACCESS.2020.2993721.

[30] S. S. Sadiq, “Improving CBIR Techniques with Deep Learning Approach: An Ensemble Method Using NASNetMobile, DenseNet121, and VGG12,” J. Robot. Control, vol. 5, no. 3, pp. 863–874, Apr. 2024, doi: 10.18196/JRC.V5I3.21805.