AI-Driven Colorectal Cancer Detection Using Machine Learning and Deep Learning

A project report submitted in partial fulfillment of the requirements for the award of the degree of

Master of Computer Applications

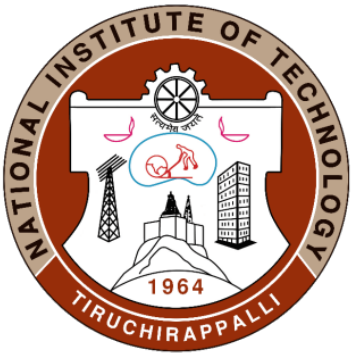
in

Computer Applications

submitted by

BIKASH KUMAR YADAV

(205121034)



Department of Computer Applications

National Institute of Technology

Tiruchirappalli, 620015

June, 2024

Candidate's Declaration

I hereby declare that the work presented in the report entitled AI-Driven Colorectal Cancer Detection Using Machine Learning and Deep Learning in partial fulfilment of the requirements for the award of the degree of **Master of Computer Applications** submitted in the Department of Computer Applications of the **National Institute of Technology Tiruchirappalli**, Tamilnadu, India is an authentic record of my own work carried out during a period from Jan 2024 to June 2024 under the supervision of **Dr.Sindhia Lingaswamy** , Department of Computer Applications, **National Institute of Technology Tiruchirappalli**. The matter presented in this report has not been submitted by me for the award of any other degree of this or any other Institute/University.

(BIKASH KUMAR YADAV)

(205121034)

Bonafide Certificate

This is to certify that the project **"AI-Driven Colorectal Cancer Detection Using Machine Learning and Deep Learning"** is a project work successfully done by **BIKASH KUMAR YADAV (205121034)** in partial fulfilment of the requirements for the award of the degree of **Master of Computer Applications** from **National Institute of Technology, Tiruchirappalli,** during the academic year 2021-2024 (6th Semester, CA750 Major Project Work).

Dr.Sindhia Lingaswamy
(Project Guide)

Prof. Michael Arock
(Head of the Department)

Project viva-voce held on 7th Jun 2024

Internal Examiner

External Examiner

Abstract

Colorectal cancer (CRC) is the third most diagnosed cancer globally and the second leading cause of cancer-related death after lung cancer. Precise histological categorization of CRC tissue is critical for diagnosis and patient management. However, the variety of tissue patterns in CRC histological images makes classification challenging.

This project applies Vision Transformers, a new class of deep-learning models in computer vision, to perform multiclass tissue classification on a publicly available CRC histology image dataset. The dataset consists of 5000 images with eight categories of tissue. We trained a classification model, specifically a Convolutional Neural Network (CNN), achieving an accuracy of 90.6

Our results outperform the original paper (81.4%) on the same dataset. Furthermore, our study highlights the opportunities of using Transformers in the histopathological image domain, showcasing their potential to improve the accuracy and efficiency of CRC tissue classification significantly. This advancement could lead to better diagnostic tools and enhanced patient management strategies in clinical practice.

Acknowledgments

We express our deep sense of gratitude to **Dr.G. Aghila**, Director, National Institute of Technology, Tiruchirappalli for giving me an opportunity to do this project. I wish to extend my sincere gratitude to **Dr.Michael Arock**, Professor and Head, Department of Computer Applications, National Institute of Technology, Tiruchirappalli for supporting and encouraging me to carry out this project work and duly evaluating my progress. I express my heartfelt gratitude to my internal guide **Dr.Sindhia Lingaswamy** , Department of Computer Applications, National Institute of Technology, Tiruchirappalli for her immense support and guidance during the study. I also thank the review panel for helping me identify the shortages of the project and helping me to overcome those shortages by guiding me. Finally, I would like to express my regards to all the faculty members of the Department of Computer Applications and others who have helped me develop this project directly or indirectly

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

Colorectal cancer (CRC) starts in the colon or rectum and is a major health issue. In Egypt, CRC is the third most common cancer in men and the fifth most common in women. Worldwide, CRC is the third most diagnosed cancer, following breast and lung cancers. It is also the second leading cause of cancer-related deaths. CRC often begins as noncancerous polyps that can take 10-20 years to turn into cancer. Adenocarcinoma, a type of cancer that comes from these polyps, makes up 96% of all CRC cases.

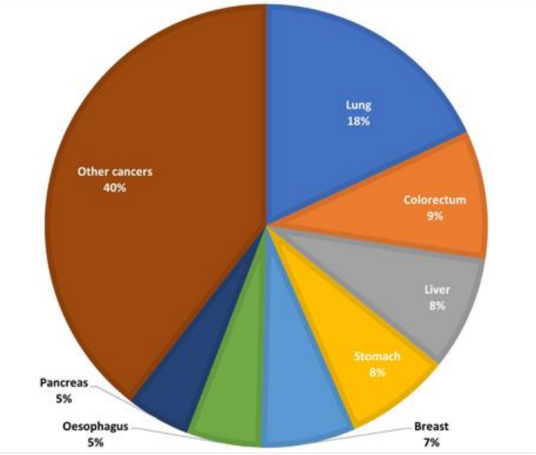


Figure 1.1: Estimated number of death cases worldwide in 2020

Human tumors, including CRC, are complex and comprise different tissue types, including tumor cells, dead tissue, immune cells, and normal tissue. Doctors use special stains to look at these tissues under a microscope. As CRC tumors change over time, their structure can provide important information about a patient's prognosis. While examining slides manually is still important, automated image processing can provide faster and more accurate analysis, which is crucial for early diagnosis and better patient care. Advances in artificial intelligence, especially deep learning, have greatly improved cancer diagnosis by automating the analysis process.

Previous methods for classifying CRC tissue images have limitations. They usually only look at two tissue types (tumor and stroma) and use private datasets, making it hard to compare results. To solve this, Kather et al. created and shared a comprehensive dataset that includes all major tissue types seen in CRC. Transformers, a new type of deep-learning model, have been successful in fields like natural language

processing and are now being applied to image analysis. This study aims to use Vision Transformers to classify different tissue types in CRC using Kather's publicly available dataset. We will compare our results with previous studies and highlight the potential of Transformers in improving histopathological image analysis. The rest of the article will discuss related work, methods, results, and conclusions.

1.1 CANCER DISEASE

Cancer, one of the most critical illnesses, originates from cells in the human body. Cells follow DNA-encoded instructions on when to divide and grow. Mistakes in DNA copying, known as mutations, can lead to uncontrolled cell growth, forming tumors. Tumors can be benign, not requiring treatment, or malignant (cancerous), spreading to other parts of the body and affecting health [14]. While specific prevention methods for cancer are lacking, risk reduction factors include quitting smoking, vaccinations, regular medical examinations, maintaining an ideal weight, regular exercise, proper nutrition, and early diagnosis [15].

1.2 COLORECTAL CANCER

Ranked among the top three most deadly cancers globally, colorectal cancer causes over 4000 deaths annually in the Kingdom of Saudi Arabia [7]. Influenced by factors like gender, age, medical condition, lifestyle choices, and genetic predisposition, early indications include low hemoglobin, changes in bowel habits, bleeding, stomach discomfort, abdominal pain, and unintended weight loss [6].

Early detection through screening, such as FOBT/FIT and colonoscopies, is crucial for recovery and improving survival rates. There are four stages of colorectal cancer, with survival probabilities varying significantly. Successful treatment may result in a survival rate exceeding 94 percent at Stage I, whereas survival rates in subsequent stages often decline [17]. Timely identification and intervention play a pivotal role in averting mortality linked to cancer [18].

1.3 ARTIFICIAL INTELLIGENCE (AI)

AI, simulating human intelligence in machines, has significantly advanced healthcare. ML, a subset of AI, focuses on enabling computers to learn and make decisions without explicit programming. DL, a

subfield of ML, uses neural networks to imitate human brain activity. AI, especially ML and DL, has applications in predicting colorectal cancer aiding in diagnosis, treatment selection, and prognosis.

1.4 MACHINE LEARNING (ML)

ML algorithms, categorized into supervised, unsupervised, and semi-supervised learning, contribute to cancer prediction. Supervised learning involves training models with labeled data, while unsupervised learning identifies patterns in unlabeled data. Semi-supervised learning combines elements of both approaches. ML, with examples like decision trees and support vector machines, offers solutions for classification and regression tasks.

1.5 DEEP LEARNING (DL)

Deep Learning, a subset of ML, focuses on hierarchical learning from experience. DL architectures include supervised, unsupervised, and hybrid networks. Convolutional Neural Networks (CNNs) are widely used in medical image analysis for cancer detection.

CHAPTER 2

LITERATURE REVIEW

Colorectal cancer (CRC) remains a significant global health challenge, underscoring the critical need for advancements in early detection methods. Researchers have explored various approaches, integrating traditional and cutting-edge technologies to enhance CRC detection accuracy and efficiency.

Previous research has highlighted the potential for AI-based techniques, including machine learning, to enhance pathologist diagnosis systems for cancer. These techniques can speed up the diagnostic process, reduce errors, and improve accuracy. The integration of deep learning systems for digital image recognition and detection marks a significant advancement in digital pathology.

2.1 TRADITIONAL SCREENING METHODS

Traditional CRC screening methods, including colonoscopy, sigmoidoscopy, and fecal occult blood tests, have been the cornerstone of detection, demonstrating effectiveness. However, these methods present challenges such as patient discomfort, compliance issues, and resource-intensive processes. Current efforts focus on integrating technological innovations to overcome these limitations, making screening more patient-friendly and resource-efficient.

2.2 IMAGING TECHNIQUES

Radiological imaging techniques, notably CT colonography and MRI, have gained prominence in colorectal cancer detection. These techniques provide detailed anatomical information crucial for identifying suspicious lesions. Challenges include the need for skilled interpretation and addressing potential radiation exposure in the case of CT colonography. Despite these challenges, the non-invasiveness and detailed imaging capabilities make these techniques valuable in the diagnostic landscape.

2.3 MOLECULAR AND GENETIC APPROACHES

Advancements in molecular and genetic research have led to exploring biomarkers and genetic signatures associated with CRC. Technologies like fecal DNA testing and blood-based assays aim to detect specific genetic alterations, offering potential non-invasive alternatives for early diagnosis.

2.4 ROLE OF MACHINE LEARNING AND DEEP LEARNING

Machine Learning (ML) and Deep Learning (DL) have emerged as transformative tools in CRC detection. ML algorithms, including support vector machines, KNN, decision tree, and logistic regression, exhibit promise in analyzing complex datasets, incorporating clinical, histopathological, and genetic information. DL models, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), excel in image recognition and temporal data analysis, significantly improving diagnostic accuracy.

2.5 RELATED WORK

In the field of CRC histology, Linder et al. (2012) and Bianconi et al. (2015) used texture-based methods to classify CRC tissue types. They extracted texture features from the images and used classification algorithms to predict tissue types. While these methods achieved high accuracies of 97-99.8%, they were limited to classifying only two tissue types (tumor and stroma), making them unsuitable for multiclass tissue classification.

In 2016, Kather et al. released a dataset of 5,000 histological images of various cancer tissue types. They evaluated multiple texture features and classification algorithms on this dataset, establishing a new benchmark with an accuracy of 85.4% for eight tissue classes. This dataset is publicly available under the Creative Commons license, encouraging further research and standardization in this field.

In 2019, Rizalputri et al. applied several multiclass classification techniques to the Kather colorectal histology dataset, including K-Nearest Neighbor (KNN), Convolutional Neural Networks (CNN), Random Forest, and Logistic Regression. Their study aimed to compare the performance of these methods, finding that CNN achieved the highest accuracy at 82.2

In 2020, Yazdi and Erfankhah proposed extracting four local features—structural, geometric, energetic, and pattern-based—to describe the diverse textures in histology images. They used the Riesz transform and monogenic local binary patterns to extract these features and tested their approach on the Kather and Kimiapath24 datasets, achieving over 88% classification and retrieval performance.

In 2021, Ohata et al. focused on the automated identification of eight tissue types in colorectal cancer histology using the Kather dataset. They employed transfer learning with CNN architectures, modifying the CNN structures to extract image features and inputting these into various machine learning algorithms.

The combination of DenseNet169 with SVM (RBF) yielded the best results, with accuracies of 87.08% and an F1-score of 88.12%.

2.6 CHALLENGES AND FUTURE DIRECTIONS

Despite promising advancements, challenges persist in the integration of these technologies into routine clinical practice. Issues such as interpretability, standardization, and ethical considerations, particularly concerning patient privacy, need careful attention. A multidisciplinary approach, combining insights from traditional and innovative methods, is pivotal for addressing these challenges and improving early detection rates, thereby enhancing patient outcomes.

The future of CRC detection hinges on a multidisciplinary approach, synergizing traditional screening methods with innovations from imaging technologies, molecular research, and artificial intelligence. Ongoing research and collaborative efforts among clinicians, researchers, and technologists will play a pivotal role in refining and implementing strategies. These endeavors aim to ultimately improve early detection rates and patient outcomes in the ongoing battle against colorectal cancer.

2.7 KEY STUDIES IN COLORECTAL CANCER DETECTION

In the pursuit of advancing colorectal cancer (CRC) detection, numerous studies have explored diverse techniques, preprocessing methods, and datasets. The following table summarizes key literature in this domain, highlighting the reference, a technique employed, preprocessing steps, the dataset used, the accuracy achieved, and notable remarks for each study.

SR No	Reference	Technique	Dataset	Accuracy	Remarks
1	Wang et al. [1]	transfer-learned deep CNN with a novel patch aggregation strategy	170,099 patches, >14,680 WSIs, >9631 subjects from China, USA, and Germany	AUC: 0.988	AI-generated heatmap highlights cancer tissue/cells.
2	Sirinukunwattana et al. [2]	Custom CNN	Utilizing more than 20,000 histology images sourced from 10 whole slide images	AUC: 0.917	Identification and categorization of nuclei
3	Xu et al. [3]	Custom Simple CNN with SVM	1376 immunohistochemistry-stained images	F1-score: 0.100, ACC: 0.100,	Categorization of epithelial and stromal regions
4	Haj-Hassan et al. [4]	Custom Simple CNN with or without initial segmentation	Hospital CHU Nancy Brabois dataset consisting of 16 multi-spectral images	ACC: 0.9917	Tumor tissue classification

Table 2.1: Literature Summary

SR No	Reference	Technique	Dataset	Accuracy	Remarks
5	Yoon et al. [5]	Utilized 6 VGG-based approaches	Examined 57 WSIs and 10,280 patches from the Center for CRC, National Cancer Center, Korea	Acc: 0.934	Binary (normal/Cancer)
6	Xu et al. [6]	Alexnet SVM	Incorporation of the CRC image dataset (1) with 717 H/E images for the 2014 MICCAI Brain Tumor Digital Pathology Challenge	ACC: 0.98 (Binary) ACC:0.872(Multiclass)	(1) Binary (cancer/not cancer (2) 6-class: normal, ADC, mucinous carcinoma,serrated carcinoma, papillary carcinoma, cribriform comedo-type adenocarcinoma
7	Shapcott et al. [7]	CNN "ciFar" model	853 images, 142 TCGA images	Detection ACC: 0.65 Classification ACC: 0.76	Analysis and identification of nuclei within the context of the research
8	Popovici et al. [8]	VGG-f (MatConvNet library)	Examined PETAC-CURACY:3 clinical trial data with 300 H/E images for a thorough investigation	Acc: 0.84	Forecasting molecular subtypes.
9	Awan et al. [9]	UNET-based architecture	sampling 139 components (71 normal, 33 low grade, 35 high grade)	Binary Acc: 0.97, 3-class ACC: 0.91	Grading of colorectal cancer

Table 2.2: Literature Summary

SR No	Reference	Technique	Dataset	Accuracy	Remarks
10	Wang et al. [10]	Architecture comprising 1 convolutional layer (1 CL) and 1 fully connected layer (1 FC), simultaneously operating on both decomposed images	University Medical Center Mannheim 1,000 images	ACC: 92.6 ± 1.2	Tumor tissue classification
11	Weis et al. [11]	VGG-based architecture incorporating residual units	HeiData Training dataset comprises 6,292 images and 20 IHC pan-cytokeratin WSIs	ACC: 0.86	Evaluation of tumor budding
12	Zhao et al. [12]	Unique architecture proposed by Ciombi, 2017	Laboratory for Pathology Eastern Netherlands 74 WSIs	ACC: 0.946	Quantification of tumor–stroma ratio (TSR) for prognosis
13	Sena et al. [13]	Custom CNN (4CL, 3FC)	Modena University Hospital, 393 WSIs	ACC: 0.817	4-class: normal mucosa, preneoplastic lesion, adenoma, cancer
14	Chuang et al. [14]	ResNet-50	Pathology Department, Chang Gung Memorial Hospital in Linkou, Taiwan, 3182 H/E WSIs	ACC: 0.999	Detection of nodal micro- and macro-metastasis

Table 2.3: Literature Summary

SR No	Reference	Technique	Dataset	Accuracy	Remarks
15	Masud et al. [15]	Custom simple CNN architecture	Utilized LC25000 dataset from James A. Haley Veterans' Hospital, consisting of 5,000 images for Colon ADC and 5,000 images for Colon Benign Tissue	ACC: .963	Binary(Colon ADC/colon benign)
16	Sarker et al. [16]	CNN architecture with weights initialized from transfer learning	14,234 images and 170,099 patches	R2 value: 0.86	Binary(cancer/not cancer)
17	Riasatian et al. [17]	Proposed a novel architecture named KimiaNet based on the DenseNet	5,000 patches	ACC: 0.963	Tumor tissue classification
18	Shimada et al. [18]	Inception.v3	Japanese cohort 201 images	ACC: 0.910	Tumor mutational burden (TMB) prediction
19	Toğaçar [19]	DarkNet-19 model based on object detection model	10,000 images	ACC: 0.999	Binary (benign/colon ADC)
20	Jiao et al. [20]	logistic regression model	180,082 patches	ACC: 0.95	Tumor tissue classification

Table 2.4: Literature Summary

CHAPTER 3

METHODOLOGY

Colorectal cancer (CRC) is a major health issue, and detecting it early and accurately is vital for effective treatment. Using medical imaging with advanced deep learning (DL) algorithms can greatly improve CRC diagnosis. This chapter explains a detailed method that uses DL techniques to analyze and interpret medical images, focusing on the important steps in this process. It will cover the dataset specifics, outline data preprocessing steps, and describe the design and implementation of the proposed model.

3.1 DATASET DESCRIPTION

In this project, we will use the colorectal cancer histology dataset provided by Kather et al., which is available online. This dataset is particularly valuable for biologists as it provides a more complex challenge than datasets like MNIST or CIFAR10, focusing on histology tiles from colorectal cancer patients. It includes eight distinct tissue classes: 'TUMOR', 'STROMA', 'COMPLEX', 'LYMPHO', 'DEBRIS', 'MUCOSA', 'ADIPOSE', and 'EMPTY'. Figure 2 illustrates a sample of these eight classes from the Kather colorectal cancer histology dataset.

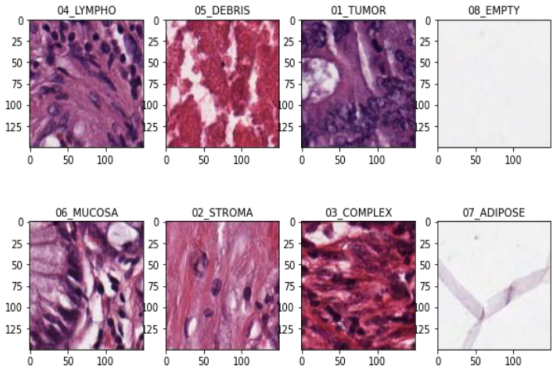


Figure 3.1: The eight classes of Kather colorectal cancer histology dataset.

The dataset comprises textures extracted from histology images of human colorectal cancer. It consists of two main archives. The first archive, "Kathertexture2016imagetiles5000.zip," contains 5000 histology images, each with a 150 x 150 pixels resolution. These images are balanced across the eight tissue types, with an equal number of images per class. The second archive, "Kathertexture2016largerimages10.zip,"

includes ten larger histology images, each measuring 5000 x 5000 pixels and combining multiple tissue types.

All images in the dataset are RGB and have a pixel size of 0.495 µm. They were scanned using an Aperio ScanScope at a 20x magnification. The histopathological specimens are anonymized images of human colorectal adenocarcinomas (primary tumors) preserved in formalin and embedded in paraffin. Each image in the dataset corresponds to one of the eight distinct tissue types.

3.2 DATA PREPROCESSING

Tuning a deep learning model's hyperparameters is an iterative process that requires a separate validation set during training and an unseen test set for final evaluation. Consequently, we divided the dataset into three parts: training, validation, and testing. We allocated 80% of the dataset for training and 20% for testing. The testing set was further split into two groups: 30% for validation during training and hyperparameter tuning, and 70% for final model evaluation to ensure no overfitting occurs. All images were normalized by standardizing the data, which helps the optimizer converge more quickly. Figure 3 displays image samples before and after standardization.

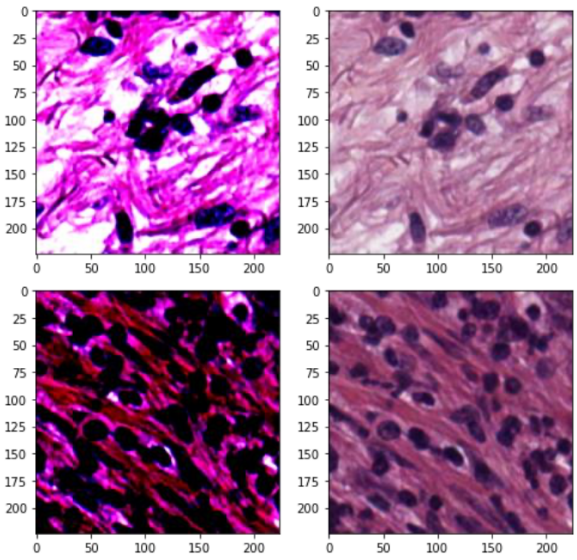


Figure 3.2: image samples before (left) and after standardization (right)

We employed various data augmentation techniques to enhance our runtime data, prevent overfitting, and enable the model to generalize well. Data augmentation is a type of regularization performed at the dataset level to mitigate overfitting and improve generalization performance by expanding the training

dataset without altering the proposed classification approach. Medical image datasets are typically small and difficult to obtain. However, Convolutional Neural Networks have shown effectiveness with data augmentation in various medical imaging tasks such as liver injury classification, brain scan analysis, and skin condition classification.

Although the dataset contains 5000 images, deep learning algorithms often perform better with more data. Data augmentation addresses the shortage of training data by making various modifications to existing data, such as scaling, flipping, and lighting adjustments, to create new images while preserving the target class's identity. This process increases the number of images available at runtime, enabling the model to train more effectively. We used various data augmentation procedures, including random translation, zooming, rotation modifications, and random horizontal flipping.

3.3 VISION TRANSFORMER

As demonstrated in the paper by Dosovitskiy et al., the Vision Transformer (ViT) architecture is based on the vanilla Transformer, which has gained significant popularity in recent years due to its high efficiency in machine translation and various natural language processing (NLP) applications. The Transformer is an encoder-decoder design that processes sequential data simultaneously without relying on a recurrent neural network. Its effectiveness is primarily due to the self-attention mechanism, which captures long-range dependencies between sequence elements.

ViT extends the application of the conventional Transformer to image classification. The goal is to generalize the Transformer's capabilities to non-textual data without incorporating any data-specific designs. ViT uses the Transformer's encoder module to perform classification by modeling a sequence of image patches into a semantic label. Unlike traditional convolutional neural networks (CNNs) that use filters with a limited receptive field, ViT's attention mechanism allows it to consider information from multiple parts of the image simultaneously, providing a more comprehensive understanding of the entire image. Figure 4 illustrates the model's complete end-to-end architecture, which generally includes:

- The embedding layer
- The encoder
- The final classifier head

The initial step involves dividing the training set images into non-overlapping patches. The Transformer treats each patch as an individual token. For an image with dimensions [c, h, w], this results in n sequence patches of dimension [c, p, p] (where c is the number of channels, h is the height, w is the width, and p is the patch size). The number of patches, n, is calculated by dividing the image's height and width by the patch size squared. Typically, a patch size of 16 or 32 is chosen because a smaller patch size results in a longer sequence, and vice versa. We selected a patch size of 15 for our image dimension of (150,150).

3.3.1 EMBEDDING LAYER

The patch sequence is linearly projected onto a 1D vector using a trainable linear projection (embedding matrix E) before being passed to the encoder. The embedded patches are combined with a learnable classification token xclass, necessary for the classification task. Positional embeddings Epos are added to the patch representations to retain positional information.The dimension of E_{pos} is $((n+1), D)$, where D is the vector size. The resultant sequence of embedded patches is represented with the token z_0, where $E \in \mathbb{R}^{(p^2 c)}$ and $E_{\text{pos}} \in \mathbb{R}^{(n+1)\times D}$, as shown in the equation:

$$z = [x_{\text{class}}; x_1 E; x_2 E; \ldots; x_n E] + E_{\text{pos}}$$

3.3.2 VISION TRANSFORMER ENCODER

The Vision Transformer (ViT) encoder consists of L identical layers. Each layer includes a multi-headed self-attention (MSA) block, represented by equation (2), and a multilayer perceptron (MLP) block, represented by equation (3). A layer normalization (LN) is applied before each block, and skip connections are used after each block. The MLP consists of two layers with a GELU non-linear activation function.

$$z'_l = \text{MSA}(\text{LN}(z_{l-1})) + z_{l-1}, \quad \text{for } l = 1 \ldots L \quad (2)$$

$$z_l = \text{MLP}(\text{LN}(z'_l)) + z'_l, \quad \text{for } l = 1 \ldots L \quad (3)$$

The final output, y, is obtained by applying layer normalization to the initial token, z_0, of the sequence from the last layer of the encoder, as shown in equation (4):

$$y = \text{LN}(z_0) \quad (4)$$

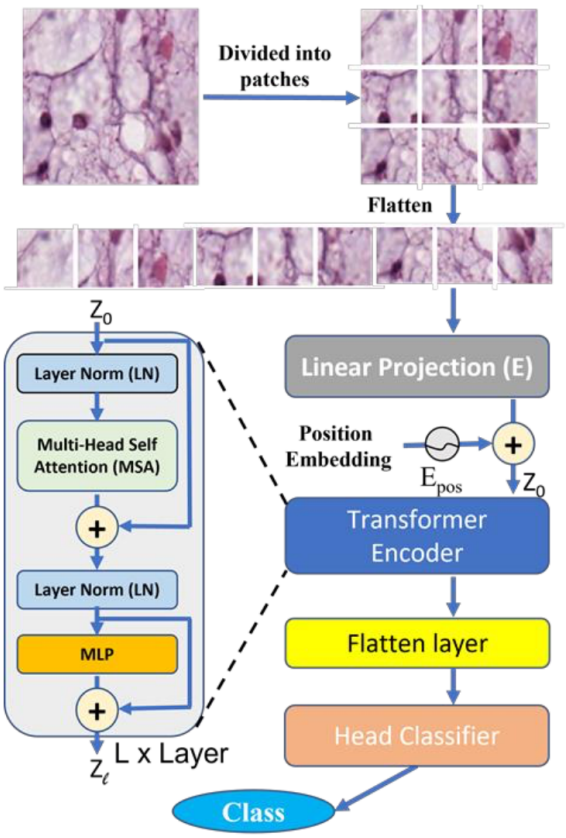


Figure 3.3: Vision Transformer (ViT) model overview.

In contrast to the original method described in the paper, where a learnable embedding classification token is prefixed to the sequence of encoded patches to serve as the image representation, we use a flattened layer to reshape the outputs of the final Transformer block. This reshaped output is used as input for the classifier head for image representation.

3.4 COMPACT CONVOLUTIONAL TRANSFORMER

Vision Transformer (ViT) architectures typically require larger datasets and longer pre-training periods due to their lack of well-informed inductive biases compared to Convolutional Neural Networks (CNNs). To address this, Hassani et al. proposed the Compact Convolutional Transformer (CCT) architecture, which integrates convolutional-based patching to retain local information and encode relationships between patches. CCTs use convolutions with short strides and introduce SeqPool, a sequential pooling approach that consolidates sequence-based information from the transformer encoder. This eliminates the need for an additional Classification Token, as shown in Fig. 5 of their study.

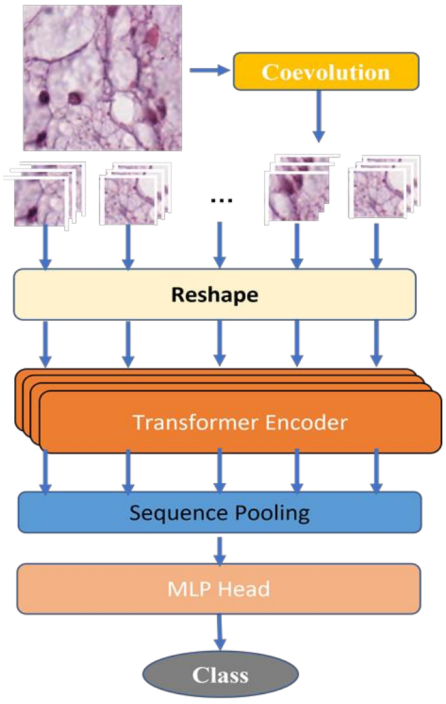


Figure 3.4: Compact Convolutional Transformer (CCT) model overview

The original study employed Auto Augment to enhance regularization for the CCT architecture. In our approach, we use simpler geometric augmentation methods like random cropping and flipping to avoid the computational cost of Auto Augment. This approach improves the generalization capability of the CCT model while streamlining the training process for practical computer vision applications.

In summary, the CCT architecture combines the strengths of transformers and convolutions, making it suitable for tasks that require both global context and local spatial features, such as image classification. This hybrid approach offers a promising solution to enhance the performance and efficiency of deep learning models in computer vision.

3.5 MODELS IMPLEMENTATION AND TRAINING

We utilized Google Colab with GPU for training. Python 3.7.12 and TensorFlow 2.6 via Keras 2.6 were used for model implementation. We employed the Adam optimizer with a learning rate of 0.001 for both ViT and CCT models. The training durations were approximately 60 minutes for ViT and 98 minutes for CCT.

Hyperparameter	**ViT Model**	**CCT Model**
Batch size	64	32
Number of epochs	100	100
Optimizer	Adam	Adam
Learning rate	0.001	0.001
Weight decay	0.0001	0.0001
Patch size	15	Not applicable
Projection dimension	128	128
Transformer layers	8	4

Table 3.1: Training Hyperparameters

CHAPTER 4

EXPERIMENTAL RESULTS

In this chapter, we will represent and explain the result of our work. We will use the accuracy, precision, recall, and f1- score to assess our model.

4.1 PERFORMANCE EVALUATION

The confusion matrix is a vital performance metric for evaluating machine learning models' classification performance. It succinctly summarizes classification outcomes with counts of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These counts serve as the foundation for computing various performance metrics. The breakdown is as follows:

True Positive (TP):	Instances correctly classified as positive.
True Negative (TN):	Instances correctly classified as negative.
False Positive (FP):	Instances incorrectly classified as positive (Type I error).
False Negative (FN):	Instances incorrectly classified as negative (Type II error).

Derived metrics include:

1. **Accuracy:** $\frac{TP+TN}{TP+TN+FP+FN}$ - Overall accuracy of the model.
2. **Precision (Positive Predictive Value):** $\frac{TP}{TP+FP}$ - Proportion of correctly predicted positive observations out of the total predicted positives.
3. **Recall (Sensitivity or True Positive Rate):** $\frac{TP}{TP+FN}$ - Proportion of correctly predicted positive observations out of all actual positives.
4. **Specificity (True Negative Rate):** $\frac{TN}{TN+FP}$ - Proportion of correctly predicted negative observations out of all actual negatives.
5. **F1 Score:** $\frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$ - Harmonic mean of precision and recall.

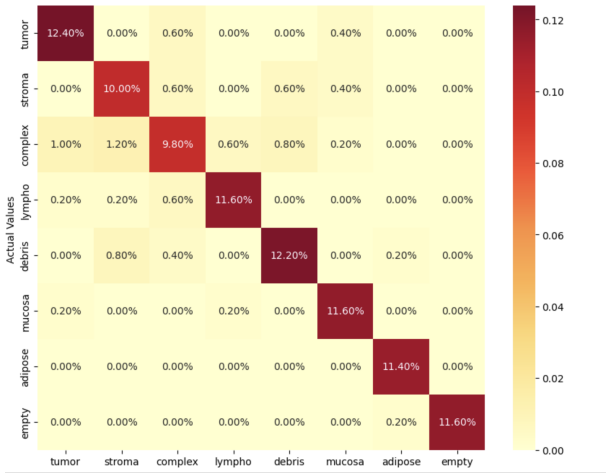


Figure 4.1: The confusion matrix for model

The ROC curve and AUC are also derived, providing insights into the model's ability to distinguish between positive and negative classes at varying threshold settings. Together, these metrics comprehensively evaluate the model's performance in classifying colorectal cancer cases.

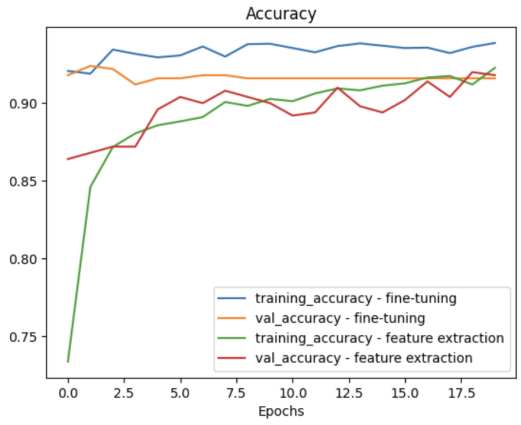


Figure 4.2: Accuracy curves Over Epochs

4.2 DEEP LEARNING MODEL PERFORMANCE

The use of convolutional neural networks (CNNs), specifically customized for colorectal cancer (CRC) detection, demonstrates the effectiveness of deep learning in the field of image-based cancer detection. The CNN model, refined by transfer learning on the CRC dataset, exhibited elevated accuracy and significant values for the area under the receiver operating characteristic (ROC) curve (AUC-ROC). This emphasizes the importance of capitalizing on pre-trained networks to enhance the model's ability to

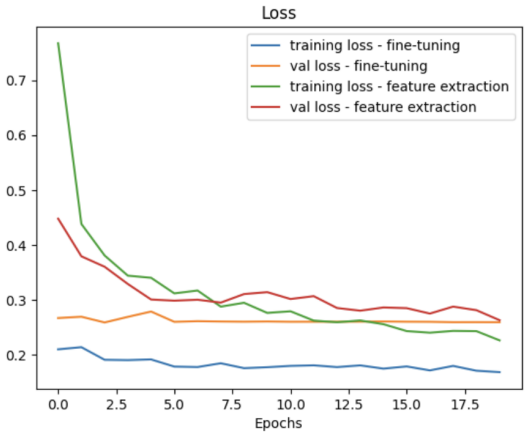


Figure 4.3: Loss curves Over Epochs

generalize to novel data.

Table 4.1: DL Performance Summary

SR No	DL Model	Dataset	Accuracy
1	CNN	5000 medical images	90.60%

```
# Evaluate the model on the test data
print("Evaluation on Test data \n")
loss, accuracy = model.evaluate(test_ds, batch_size=32)
print(f"\nModel loss on test set: {(loss):.2f}")
print(f"Model accuracy on test set: {(accuracy*100):.2f}%")

Evaluation on Test data

16/16 [==============================] - 1s 53ms/step - loss: 0.2630 - accuracy: 0.9060

Model loss on test set: 0.26
Model accuracy on test set: 90.60%
```

Figure 4.4: Performance Measures of models on testing set

Chapter 5

CONCLUSIONS

In this study, we have undertaken a comprehensive examination of 20 recent studies about utilizing machine learning (ML) and deep learning (DL) in diagnosing and prognosis colorectal cancer. To enhance the clarity of our review, we meticulously gathered these studies. We categorized them into three principal sections: the purpose of prediction, methods employed, and datasets used for prediction. We furnish a concise overview of the research within each category, presenting diverse perspectives. Our systematic organization of the studies in a tabular format facilitates a nuanced and detailed comparison.

Our findings underscore a recent research emphasis on developing predictive models utilizing DL techniques, primarily aimed at forecasting normal or abnormal conditions using publicly available datasets. Furthermore, we offer an in-depth analysis of the research to elucidate medical facets while delving into challenges and potential advancements in DL applications for colorectal cancer prediction.

In conclusion, the significant impact of Artificial Intelligence (AI) on healthcare is evident, particularly in improving the prediction of malignancies such as colorectal cancer. The application of DL algorithms signifies a noteworthy advancement in healthcare practices.

REFERENCES

[1] K.-S. Wang, G. Yu, C. Xu, X.-H. Meng, J. Zhou, C. Zheng, Z. Deng, L. Shang, R. Liu, S. Su *et al.*, "Accurate diagnosis of colorectal cancer based on histopathology images using artificial intelligence," *BMC medicine*, vol. 19, pp. 1–12, 2021.

[2] K. Sirinukunwattana, S. E. A. Raza, Y.-W. Tsang, D. R. Snead, I. A. Cree, and N. M. Rajpoot, "Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1196–1206, 2016.

[3] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," *Neurocomputing*, vol. 191, pp. 214–223, 2016.

[4] H. Haj-Hassan, A. Chaddad, Y. Harkouss, C. Desrosiers, M. Toews, and C. Tanougast, "Classifications of multispectral colorectal cancer tissues using convolution neural network," *Journal of pathology informatics*, vol. 8, no. 1, p. 1, 2017.

[5] H. Yoon, J. Lee, J. E. Oh, H. R. Kim, S. Lee, H. J. Chang, and D. K. Sohn, "Tumor identification in colorectal histology images using a convolutional neural network," *Journal of digital imaging*, vol. 32, pp. 131–140, 2019.

[6] Y. Xu, Z. Jia, L.-B. Wang, Y. Ai, F. Zhang, M. Lai, and E. I.-C. Chang, "Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features," *BMC bioinformatics*, vol. 18, pp. 1–17, 2017.

[7] M. Shapcott, K. J. Hewitt, and N. Rajpoot, "Deep learning with sampling in colon cancer histology," *Frontiers in bioengineering and biotechnology*, vol. 7, p. 52, 2019.

[8] V. Popovici, E. Budinská, L. Dušek, M. Kozubek, and F. Bosman, "Image-based surrogate biomarkers for molecular subtypes of colorectal cancer," *Bioinformatics*, vol. 33, no. 13, pp. 2002–2009, 2017.

[9] R. Awan, K. Sirinukunwattana, D. Epstein, S. Jefferyes, U. Qidwai, Z. Aftab, I. Mujeeb, D. Snead, and N. Rajpoot, "Glandular morphometrics for objective grading of colorectal adenocarcinoma histology images," *Scientific reports*, vol. 7, no. 1, p. 16852, 2017.

[10] C. Wang, J. Shi, Q. Zhang, and S. Ying, "Histopathological image classification with bilinear convolutional neural networks," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2017, pp. 4050–4053.

[11] C.-A. Weis, J. N. Kather, S. Melchers, H. Al-Ahmdi, M. J. Pollheimer, C. Langner, and T. Gaiser, "Automatic evaluation of tumor budding in immunohistochemically stained colorectal carcinomas and correlation to clinical outcome," *Diagnostic pathology*, vol. 13, pp. 1–12, 2018.

[12] K. Zhao, Z. Li, S. Yao, Y. Wang, X. Wu, Z. Xu, L. Wu, Y. Huang, C. Liang, and Z. Liu, "Artificial intelligence quantified tumour-stroma ratio is an independent predictor for overall survival in resectable colorectal cancer," *EBioMedicine*, vol. 61, 2020.

[13] P. Sena, R. Fioresi, F. Faglioni, L. Losi, G. Faglioni, and L. Roncucci, "Deep learning techniques for detecting preneoplastic and neoplastic lesions in human colorectal histological images," *Oncology Letters*, vol. 18, no. 6, pp. 6101–6107, 2019.

[14] W.-Y. Chuang, C.-C. Chen, W.-H. Yu, C.-J. Yeh, S.-H. Chang, S.-H. Ueng, T.-H. Wang, C. Hsueh, C.-F. Kuo, and C.-Y. Yeh, "Identification of nodal micrometastasis in colorectal cancer using deep learning on annotation-free whole-slide images," *Modern Pathology*, vol. 34, no. 10, pp. 1901–1911, 2021.

[15] M. Masud, N. Sikder, A.-A. Nahid, A. K. Bairagi, and M. A. AlZain, "A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework," *Sensors*, vol. 21, no. 3, p. 748, 2021.

[16] M. M. K. Sarker, Y. Makhlouf, S. G. Craig, M. P. Humphries, M. Loughrey, J. A. James, M. Salto-Tellez, P. O'Reilly, and P. Maxwell, "A means of assessing deep learning-based detection of icos protein expression in colon cancer," *Cancers*, vol. 13, no. 15, p. 3825, 2021.

[17] A. Riasatian, M. Babaie, D. Maleki, S. Kalra, M. Valipour, S. Hemati, M. Zaveri, A. Safarpoor, S. Shafiei, M. Afshari *et al.*, "Fine-tuning and training of densenet for histopathology image representation using tcga diagnostic slides," *Medical Image Analysis*, vol. 70, p. 102032, 2021.

[18] Y. Shimada, S. Okuda, Y. Watanabe, Y. Tajima, M. Nagahashi, H. Ichikawa, M. Nakano, J. Sakata, Y. Takii, T. Kawasaki *et al.*, "Histopathological characteristics and artificial intelligence for predicting tumor mutational burden-high colorectal cancer," *Journal of gastroenterology*, vol. 56, no. 6, pp. 547–559, 2021.

[19] M. Toğaçar, "Disease type detection in lung and colon cancer images using the complement approach of inefficient sets," *Computers in Biology and Medicine*, vol. 137, p. 104827, 2021.

[20] Y. Jiao, J. Yuan, Y. Qiang, and S. Fei, "Deep embeddings and logistic regression for rapid active learning in histopathological images," *Computer Methods and Programs in Biomedicine*, vol. 212, p. 106464, 2021.