Proceedings of the 5th International Conference on Advanced Science & Engineering ICOASE 2025, University of Zakho and Duhok Polytechnic University, 24th – 25th September 2025

(Special issue for Passer Journal of Basic and Applied Sciences)
Paper No. 15

Detection and Classification of Colorectal Cancer Types Using Deep Residual Learning based on ResNet-50 with Adam Optimization Method

Herman Khalid Omer [1], Salwa M. Hasan [2], Lozan M. Abdullrahman [1], Amira Bibo Sallow[1]

[1] IT Department, Duhok College Technique, Duhok Polytechnic University, Duhok 42001, Kurdistan Region, Iraq.
[2] Food Production and Technology Department, Duhok College Technique, Duhok Polytechnic University, Duhok 42001, Kurdistan Region, Iraq.
Email: herman.omer@dpu.edu.krd , salwa.muhsin@dpu.edu.krd , lozan.abdullrahman@dpu.edu.krd, amira.bibo@dpu.edu.krd

Abstract

These days, digital pathology plays a crucial role in tumor detection and prognosis, particularly in objectively evaluating histological images for colorectal cancer (CRC). Deep learning models have demonstrated significant success in image classification, prompting researchers to adopt these methods for medical image analysis. This research explores deep learning methods, specifically using ResNet architectures (ResNet-18, ResNet-50, ResNet-101), combined with various optimization methods, including Adam, Stochastic Gradient Descent with Momentum (SGDM), and Root Mean Square Propagation (RMSProp), for classifying colorectal cancer types from histological images. A comprehensive evaluation was conducted to identify the most effective neural network architecture and optimization approach. Among the tested combinations, the ResNet-50 model with the Adam optimization method achieved the highest accuracy of 99.86% on the NCT-CRC-HE-7K dataset and an accuracy of 98.36% on the NCT-CRC-HE-100K dataset, outperforming other tested combinations and existing approaches in the literature, effectively differentiating between benign and aggressive colorectal cancer. Lastly, the suggested model was compared to the current models, which were assessed to ascertain the most effective neural network models and the best training approach for our case study on colon tumor segmentation.

Keywords: Colorectal Cancer; Deep Residual Learning; ResNet-50; Adam Optimization.

1 Introduction

Globally, Colorectal Cancer (CRC) is the second leading cause of cancer-related mortality and the third most prevalent kind of cancer diagnosed [1]. These days, there are more cancer patients worldwide, and the disease's prevalence is rising annually. Personal characteristics or behaviors including age, history of chronic illness, polyps, genetics, and lifestyle are linked to the development of cancer [2].

While intestine mucosal polyps are frequently the initial sign of colorectal cancer, it can also begin as an adenoma, a benign lesion that, depending on its size and histological appearance, can develop into a malignant lung cancer [3]. Hematoxylin-eosin (HE)-stained tissue slides are available for almost all CRC patients [4]. Several common tissue types, including the primary tissue stains used in histology, polyps, Cancer-Associated Stroma (STR), Adipose Tissue (ADI), Normal Colon Mucosa (NORM), and Lymphocytes (LYM) [5].

The diagnosis and the ability to predict the severity of the disease, in order to apply the most effective treatment, depend on the early and accurate detection of colorectal cancer. [18,6]. There are two types of colorectal examinations: complete colonoscopies and the Fecal Immunological Test (FIT) or Fecal Occult Blood Test (FOBT) [7] Colonoscopy is the gold standard for investigating CRC. To identify patients who are most at risk of developing CRC and to shorten the time between the initial consultation and the colonoscopy, a number of risk classification scores based on symptoms have been developed [8,9,10]. Numerous research' findings have demonstrated that colorectal cancer growth may be accurately predicted by classifying medical images [4,5].

Over the past several years, Artificial Intelligence (AI) has gradually made its way into the medical and pathological fields [11]. Machine Learning (ML) is one of the subfields of AI. Its main objective is to analyze and understand data structures and patterns so that learning, inference, and decision-making may occur without the need for human

intervention. As a branch of ML, basically, Deep Learning (DL) is a neural network that has three or more hidden layers. Additionally, it has been demonstrated that DL has yielded superior results in a number of domains, including as medical imaging, identification of object, and speech recognition [13,14]. This study therefore sought to predict colorectal cancer using deep learning algorithms based on images from histological images of human CRC and normal tissue that were discolored with hematoxylin and eosin (H&E), as well as to predict the patients' normal or abnormal condition.

CRC is one of the major causes of cancer-related mortality, hence early detection is crucial. Traditional diagnosis is time-consuming and subjective and hence requires automated deep learning-based approaches. ResNet-50 combined with Adam optimization is used in this paper to enhance classification accuracy.

Our main contributions are summarized as follows:

• We present a controlled comparative study of multiple ResNet backbones (ResNet-18, ResNet-34, ResNet-50, ResNet-101) for colorectal cancer histopathology image classification under a unified preprocessing and training pipeline. All models were trained and evaluated using the same dataset, same augmentation steps (median filtering, sharpening, resizing), and same protocol to ensure a fair comparison.

• We systematically analyze the effect of different optimizers (Adam, Stochastic Gradient Descent with Momentum (SGDM), and RMSProp) on each backbone and show that the pairing of ResNet-50 with Adam achieves the most stable and reliable performance, including very high recall for malignant tissue classes. This joint backbone–optimizer analysis is not emphasized in prior colorectal cancer classification studies, which typically report only a single "best model."

• We provide detailed evaluation using accuracy, precision, recall, F1-score, confusion matrix, and ROC/AUC, and we include cross-validation to demonstrate that the performance is consistent and not dependent on a single split. This directly addresses robustness and reduces concerns of overfitting on histopathology patches.

• We position the highest-performing configuration (ResNet-50 + Adam) as a practical triage support tool in digital pathology. Rather than replacing the pathologist, the model is intended to highlight suspicious colorectal regions with low false negatives, helping the expert prioritize areas of interest during slide review. This differs from prior work such as Kather et al., which focuses on prognosis and survival stratification, and from general CNN benchmarking studies, by targeting immediate assistive use in routine diagnostic workflows.

The remainder of the paper follows this format: In Section 2, relevant research on colorectal cancer diagnosis using deep learning is presented. The dataset, preprocessing, and implementation of the ResNet-50 are all explained in Section 3 along with the materials and procedures. Results are presented along with an error analysis in Section 4. Key findings and discusses future research directions concluded in Section 5.

2 Related Work

This section highlights significant advancements in the diagnosis of colorectal cancer across three main areas: histopathological image analysis, radiological imaging techniques, and the integration of genomic and biomarker data. It explores current research trends and recent innovations in each domain. In addition, the role of IoT-enabled devices, sensor-based systems, and automated diagnostic frameworks is emphasized, particularly in supporting early detection, precise tumor localization, and personalized treatment planning. The reviewed studies demonstrate how deep learning and machine learning models have been increasingly applied to improve diagnostic accuracy, streamline data processing, and address workflow bottlenecks within clinical environments.

Kumar, V.T.R.P. et al. [15] Conducted a systematic comparative analysis of five Deep Learning models—Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Transfer Learning, AlexNet, and GoogLeNet—for classifying colon cancer using CT colonography images. The proposed framework involved three stages: pre-processing (noise removal via median filtering), segmentation (affected region isolation using SegNet), and classification. The dataset consists of 1,000 images (700 training, 300 testing) from 825 clinical cases, with polyps of all sizes and some negative cases. The workflow was thus geared towards rigorous feature extraction and computational efficiency, with GoogLeNet topping the list. With 94.165% accuracy, 97.589% sensitivity, and 87.359% specificity for 60% training data and 5-fold cross-validation, it outperformed other models owing to onset modules and multi-scale feature fusion. The study used a small dataset, limiting generalization, and focused only on CT colonography images, which may not be ideal for early-stage colorectal cancer detection.

Sirinukunwattana, K. et al. [16] provided a deep learning framework that performs image-based consensus molecular subtype (imCMS) classification, directly predicting the colorectal cancer (CMS) subtypes from routine hematoxylin and eosin (H&E)-stained histology slides. The technique depends upon Inception V3 neural networks trained using three independent cohorts (FOCUS, TCGA, GRAMPIAN; total: 1,540 histology slides) to associate histologic patterns to transcriptomically defined CMS groups (CMS1-4). By employing domain adversarial training to alleviate dataset variability, imCMS performed robustly (macro-average Area Under the Curve AUC: 0.88-0.85) and resolves intratumoral heterogeneity. The imCMS framework fuses histopathology with molecular oncology enabling H&E based CMS classification with an accuracy of more than 85%. The model is solely built upon Inception V3, without the possible performance benefits from other architectures like ResNet, and relies on multi-omic datasets, which are not always available within clinical settings.

On the basis of the work presented in Bokhorst, J.-M. et al. [17] where deep learning-based multiclass semantic segmentation is applied, this method automates segmentation and modeling of 14 tissue compartments for CRC histopathology images. The model is designed having in mind the difficulty faced by CRC diagnosis in segmentation as it is trained from diverse multi-centric public datasets. Multiple loss functions (categorical cross-entropy, Focal, Bi-tempered, and Lovasz-softmax) have been further tested, with Lovasz-softmax displaying the highest performance. The segmentation model has been embedded in a Computer-Aided Diagnosis (CAD) system to classify the biopsies into four clinically relevant categories (high-risk, low-grade dysplasia, hyperplasia, and benign). Their validation on an independent cohort of over 1,000 patients was notable in demonstrating strong diagnostic precision with great potential in pathological workflow streamlining. The study focused primarily on segmentation rather than classification, limiting its applicability to automated diagnostic decision-making.

Tsai, M.-J. et al. [18] examined the use of deep learning techniques, and particularly convolutional neural networks (CNNs), for the classification of colorectal cancer (CRC) histopathology. The authors presented a systematic comparison of the performances of five different CNN architectures (i.e., AlexNet, SqueezeNet, VGG19, GoogLeNet, ResNet50) and three optimizers (SGDM, RMSProp, Adam) to identify the best model for tissue classification. This exploration made use of two datasets, namely NCT-CRC-HE-100K (nine tissue classes, 100,000 images) and Kather-texture-2016-image (eight tissue classes, 5,000 images). For this project, the ResNet50 + RMSProp model garnered an astounding 99.69% accuracy for the NCT-CRC-HE-100K dataset (9-class) and 94.86% for the Kather-texture-2016-image (8-class) surpassing the previous techniques (87.4% for 8-class, etc.). It was also an improvement over VGG19 (99.69% vs. 94.3% on external validation) and upheld robustness in the face of staining/scanner variability. Tumor and stroma classes demonstrated >97% precision; the training time was ~40 minutes per epoch on an NVIDIA GTX 1070 GPU. While the study provided a comprehensive performance comparison, it lacked an in-depth evaluation of optimization methods, particularly Adam, which has been shown to enhance learning stability.

Jiang's [19] introduced an open-source, Colorectal cancer prognosis method based on deep learning that uses self-supervised learning and attention-based multiple-instance learning. The model, trained on 908 patients from Germany, achieved significant prognostic accuracy, achieving C-indexes of 0.76–0.82 for survival prediction and hazard ratios up to 8.35 for high-risk patients. The study focused on prognosis rather than classification, leaving a gap in automated early detection.

Prezja's [20] introduced a refined deep learning approach to enhance tissue decomposition accuracy in colorectal cancer histopathology images, achieving 95.6% accuracy on external and 99.5% internal testing, reducing errors in biomarker-relevant classes. The approach primarily targeted biomarker-relevant classification rather than holistic cancer detection.

Masud, M. et al. [21] suggested a deep learning-based system to diagnose cancer by classifying histological images of colon and lung tissues into five groups (three malignant, two benign). Training on the LC25000 histopathological image dataset which is a dataset composed of 25,000 histopathological images, a multi-channel CNN architecture model was able to achieve a 96.33% accuracy rate. The framework that combines Fourier and wavelet stands out both in terms of a more diverse set of features and classification reliabilities, which leads the model to offer a cost-effective solution for early cancer detection in low-resource settings. The study used a dataset combining lung and colon cancer images, which may have affected specificity in distinguishing between different cancer types.

A researcher, Nemlander, E. et al [22], recently provided findings from a study they conducted developing a machine learning approach to detect NMCRC cases in primary care patients. The research was also aimed at identifying symptoms and improving the patterns of diagnosis and consulting to detect the disease in its early stages. In applying stochastic gradient boosting to derive a model in the isolated NMCRC from 542 cases, 2,139 matched controls from

Sweden, it identified symptoms like anemia, alteration of bowel habit, and rectal bleeding as indicators. The study finds a sensitivity of 73.3% and a specificity of 83.5% and demonstrates the capacity of machine learning to develop early cancer detection in primary care. The model relied heavily on clinical symptoms rather than histological image-based diagnostics, limiting its applicability to automated classification systems.

Using RNA-seq data from extracellular vesicles in machine learning and deep learning models, Bostanci et al. [23] predicted the presence of colon cancer and used data to stage the disease from a total of 150 subjects participating in the study, 50 of whom were healthy and 100 were suffering from cancer stages I-IV. These data too loud into 300 records using SMOTE technique for balancing and were min-max normalized. Feature selection was performed based on Information Gain (IG) to target the miRNA transcripts (e.g., hsa-miR-335-5p) as primary biomarkers. Five canonical Machine Larning (ML) models (K-Nearest Neighbors (KNN), Logistic Model Tree (LMT), Regression Tree (RT), Random Classification (RC), Random Forest (RF)) along with three DL architectures (1D-CNN, LSTM, BiLSTM) were tested, where the DL hyperparameters were optimized using a GA. 1D-CNN had a predictive accuracy of 97.67% for cancer prediction while Bidirectional Long Short-Term Memory (BiLSTM) achieved accuracy of 98% in stage classification, surpassing the ML models. Feature selection boosted ML performance (e.g., 97.33% accuracy with RF using 30–40 identified features), while DL models performed better when using full feature sets. The method requires Ribonucleic Acid Sequence (RNA-seq) data, which is less useful for everyday diagnosis since it is less accessible than histopathological pictures.

Previous research centers on segmentation, prognosis, or molecular subtypes rather than on the optimization of classification accuracy. Beyond being based on limited datasets or dedicated imaging modalities, they lack comprehensive assessments of deep learning models. In order to fill these gaps, our work compares ResNet models with optimizers, obtaining excellent accuracy for accurate classification of colorectal cancer.

3 Methods and Materials

This sector provides a full explanation of the materials and methods used in proposed colon cancer categorization strategy. The dataset and preparation methods utilized in this work are also described, along with the approaches employed to increase the data's quality and appropriateness for neural network analysis. An overview of the model's workflow of proposed transfer learning system is shown in Figure 1.

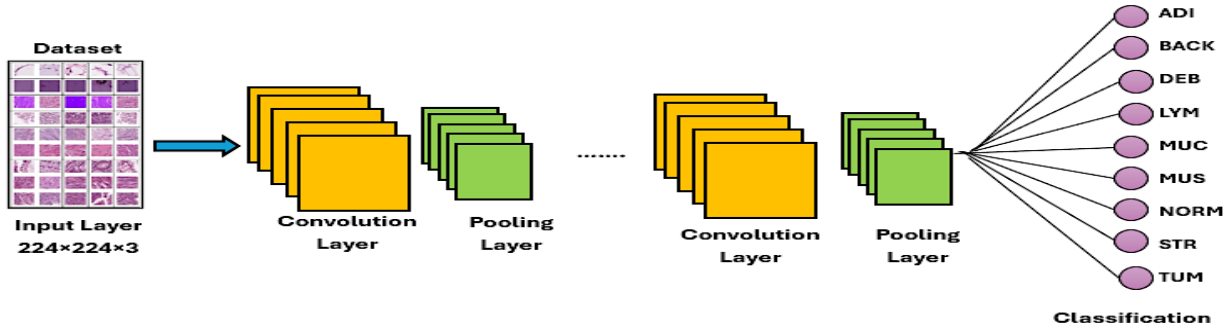


Figure 1: Proposed Classification Colon Cancer Diagram Using Transfer Learning.

3. 1 Dataset Description

The NCT-CRC-HE-7K dataset was used to validate the suggested models after they were trained on the NCT-CRC-HE-7K dataset, which was consisted of 7180 image patches from fifty patients with colorectal adenocarcinoma. The collection comprises 86 H&E-stained human cancer tissue slides with 224x224 × 3 pixels at 0.5 microns per pixel. Images of initial tumor slides and liver metastases are included, and they have been color-normalized using Macenko's approach. Figure 2 shown nine Samples of each tissue classes of NCT-CRC-HE-7K, and Figure 3 shown number of images for each class from NCT-CRC-HE-7K.

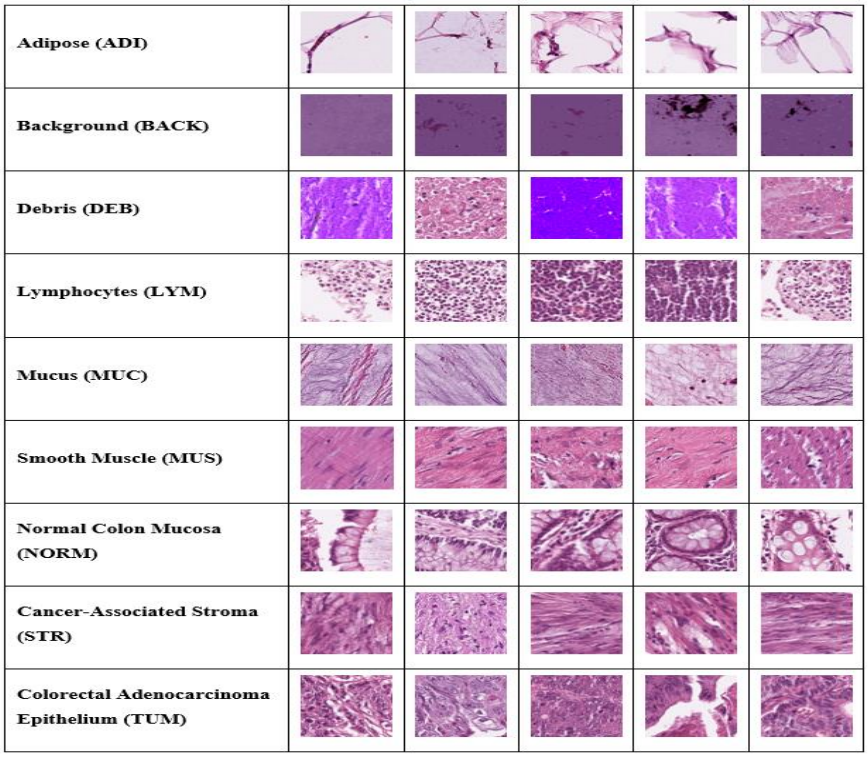


Figure 2: Nine Samples of each Tissue classes of NCT-CRC-HE-7K Dataset [24].

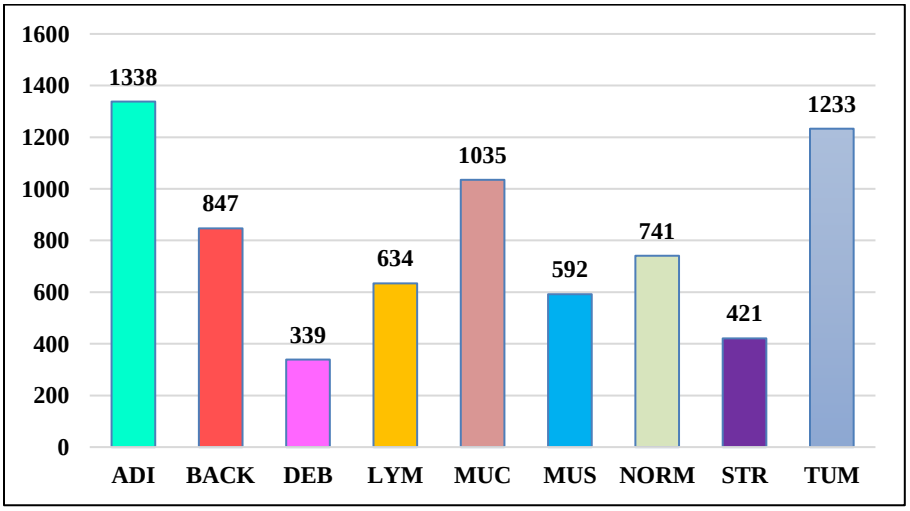


Figure 3: Number of images for each class from NCT-CRC-HE-7K Dataset

3. 2 Image Preprocessing

Preprocessing images is critical to image processing techniques as this converts underdone data into a format for easier analysis and enhancement of the most important image qualities such as contrast, resolution, and noise level. This is important in all the applications of computer vision and machine learning, especially in the field of medical imaging. In order to achieve extremely precise results, the proposed model integrates three essential preprocessing methods: Median filtering, sharpening filtering and image reshaping. By reduce noise using Median Filter for and sharping filter for whole the dataset images, as shown in Figure 4.

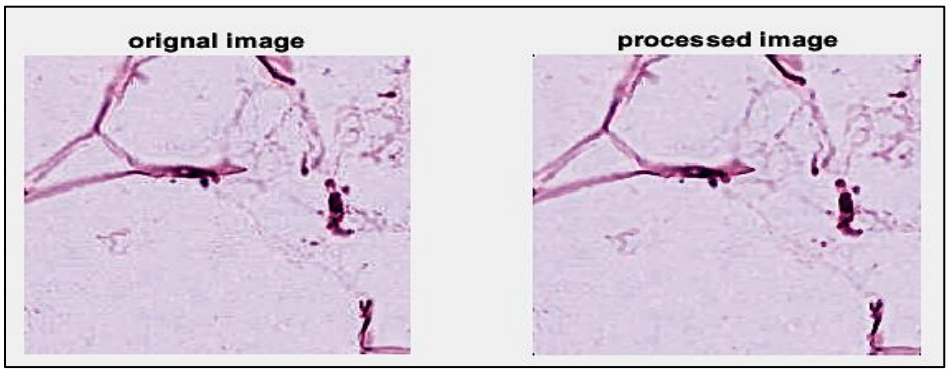


Figure 4: Preprocessed Result on Input Image.

3. 3 ResNet in Deep Residual Learning

Modern technology has brought about a fantastic evolution in computer vision, which uses deep models to solve diverse problems, such as image recognition, object detection, face recognition, and image classification. Today deep Convolutional Neural Networks (CNNs) are being proposed for accurate analysis of visual imagery. Nonetheless, increasing the number of layers in CNNs can make the training ever more difficult while producing reduced accuracy. ResNet helps to overcome some of these challenges with regard to training the deep neural networks. ResNet is a deep neural network that uses residual learning to make training very deep networks feasible. It was introduced in a paper by He, Zhang [25], Ren, and Sun, and is known for its outstanding performance in image recognition tasks. By including skip connections, ResNet enables the model to learn the distinction between input and the intended output instead of learning the output directly. Mathematically, a residual block can be expressed as:

$$Output = F(Input, \{W_i\}) + Input \quad (1)$$

Where:

- $F(Input, \{W_i\})$: This is the transformation over the input (convolution, activation, etc.).
- Introducing the input allows the network to learn the residual, which is the difference of the input and output, and furthermore, enables easier training.

The ResNet Architecture has been shown in general in Figure 5, and ResNet Architecture variants:

- ResNet-18: a ResNet containing 18 layers.
- ResNet-34: contains 34 layers of a ResNet.
- ResNet-50: it is a form of ResNet containing 50 layers. This uses bottleneck layers for greater efficiency in organization as well.
- ResNet-101, ResNet-110, ResNet-152: more complicated versions of ResNet that can contain hundreds of layers.

The Resnet50 architecture, based on the 50-layer ResNet model, features a bottleneck design with a stack of 3 layers, resulting in higher accuracy and a performance of 3.8 billion FLOPs, compared to the earlier 2-layer Resnet model:

i. A preprocessing layer that typically includes a convolution, batch normalization, and Rectified Linear Unit (ReLU) activation.

ii. A series of residual blocks that may include multiple stages with different numbers of filters.

iii. A fully connected (dense) layer at the end that produces the final output (such as for classification).

iv. The final layer usually uses *softmax* or another activation function, depending on the task.

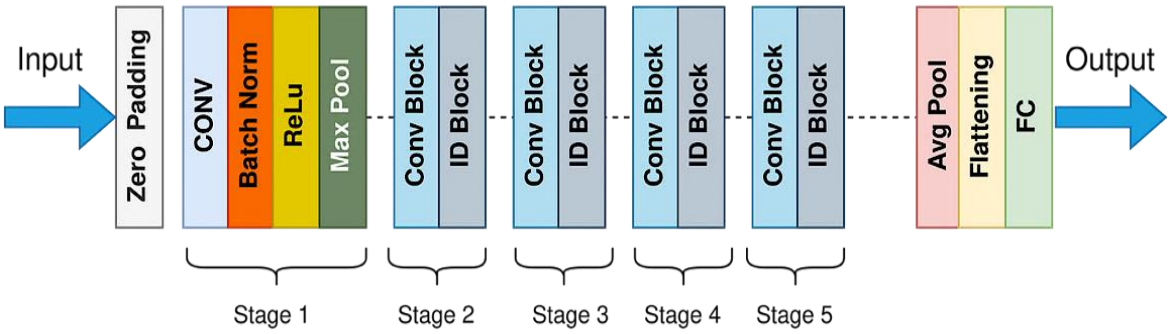


Figure 5: ResNet Model Architecture

Accuracy, Precision, and Recall are the metrics used in performance evaluation calculations for the experiment findings. The following is implied by the accuracy, precision, and recall theorems:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

where *TP* is the number of true positives, *TN* is the number of true negatives, *FP* is the number of false positives, and *FN* is the number of false negatives.

3. 4 Optimization in Deep Learning

The optimizers for deep learning, such as Stochastic Gradient Descent (SGD), Adam, and RMSprop, tune the model parameters to reduce the loss function. They aid in the learning of the neural networks by updating the weights and biases according to the feedback on performance on the data. These algorithms improve performance by optimizing the parameters of the model.

- Use the adam (adaptive movement estimation) algorithm. It can adjust the gradient decay rates and squared gradient moving averages using the Optimizer Parameters option.
- Use the Stochastic Gradient Descent with Momentum (SGDM) algorithm. it can specify the momentum value using the Momentum field of the Optimizer Parameters option.
- Use the Root Mean Square Propagation (RMSProp) algorithm. The squared gradient moving average's decay rate can be adjusted using the SGDM Factor fields in the Optimizer Parameters option.

Adam optimizes the learning models through a mechanism in which the learning process is adaptive to the context, a situation that applies well in training machine learning models, reinforcements models like those of the actor-critic. Adam is the generalization of two other variants of SGDM and RMSProp. This is how Adam works:

- Momentum helps accelerate gradient descent in the relevant direction, dampening oscillations.
- RMSProp uses the average of recent squared gradients to modify the learning rate for each parameter.

Adam uses gradient moving averages and squared gradients to update model parameters, enhancing stability and smoothing the learning process:

- Learning Rate (α): The step size used in the update rule.
- β_1: The decay rate for the first moment (gradient moving average). This is often close to 1, e.g., 0.9.

- β_2: The decay rate for the second moment (squared gradient moving average). This is also close to 1, e.g., 0.999.
- ε: A little constant that is added to the denominator in order to prevent division by zero (typically 1e-8).

The moving averages are updated using the following equations:

a. First Moment Estimate (m^t):

$$m^t = \beta_1 \cdot m_{-1}^t + (1 - \beta_1) \cdot g^t \quad (5)$$

where g^t is the gradient at time step t.

b. Second Moment Estimate (v^t):

$$v^t = \beta_2 \cdot v_{-1}^t + (1 - \beta_2) \cdot g^{t2} \quad (6)$$

c. Bias-corrected First and Second Moment Estimates: These are used to counteract the initialization bias of m^t and v^t (i.e., the fact that the estimates start at zero):

$$\widehat{m}^t = \frac{m^t}{1-\beta_1^t}, \quad (7)$$

$$\hat{v}^t = \frac{v^t}{1-\beta_2^t} \quad (8)$$

d. Parameter Update:

$$\theta^t = \theta_{-1}^t - \frac{\alpha \cdot \widehat{m}^t}{\sqrt{\hat{v}^t}+\varepsilon} \quad (9)$$

where θ^trepresents the model parameters at step *t*.

3. 5 Experimental Setup and Reproducibility

The experiments were implemented in a fixed computing environment. The system specifications (hardware and software) are listed in Table 1 to support reproducibility.

Table 1: System specifications used for all experiments.

Specification	Detail
Operating System	Windows 11 Home
CPU	Intel Core-i7 10th Gen.
RAM	32 GB
GPU	NVIDIA RTX 2070 (8 GB vRAM)
Software Version	MATLAB 2024b

4 Results and Discussions

In this study, A distribution of 90% training data and 10% testing data were used in the studies. The experiment's early phases, the preprocessing step is done using median filter for noise reduce and sharpen filter to emphasize regions with high spatial frequency in order to highlight details. The next step created proposed ResNet-50 model with suitable size of color image from dataset, and used variable as shown in Table 2.

Table 1: Properties of ResNet-50 model.

Variable	Value
Image dimensions	224 × 224 × 3
Channels	9
Epochs	10
Batch size	32
Initial Learn Rate	10^{-4}
2DConvolution layers (size)	3 (3 × 3)
Convolution layer activation	ReLu
2DMaxpooling layers (size)	3 (3 × 3)
Validation Frequency	30
Dense layer activation	*softmax*
Compiler optimizer	adam
Compiler loss	Categorical *crossentropy*

Every ResNet-50 architectural layer has certain requirements for creation. The convolution layer is typically used to create the feature learning model. The ResNet-50 design consists of a convolution layer, activation layer, pooling layer, and batch normalization layer. After the convolution layer, batch normalization is used, and the ReLu activation function is used in the activation layer. The convolution layer in each ResNet-50 design is divided into five layers by residual blocks. Only at the start of feature learning, following the first convolution, and at the end of feature learning, following the preceding convolution, are pooling layers used before being incorporated into the classification layer. After that, the system is going through the preprocessing stage. Table 2 is a comparative evaluation of accuracy results using three ResNet architectures (ResNet-18, ResNet-50, ResNet-101) and three optimization algorithms: Adam, SGDM, and RMSProp. Out of the architectures used, ResNet-50 achieved better accuracy rates than ResNet-18 and ResNet-101. Specifically, with Adam optimizer, ResNet-50 had the highest accuracy, which was 99.8% at training and 99.4% on testing. In comparison, ResNet-18 was less accurate (99.1% training and 98.4% testing), and ResNet-101 was of intermediate accuracy (99.3% training and 99.0% testing).

In terms of optimization methods, Adam overall had superior accuracy for all the ResNet models, with SGDM a close second, and RMSProp having the worst performance of the tested optimizers. The results of these results indicate that the adaptive learning capability of the Adam optimizer, combined with ResNet-50 architecture, produces the optimal means of accurately diagnosing colorectal cancers from histologic images. This comparison confirms the validity of applying ResNet-50 with Adam optimization as our optimal selection in this research.

Table 2: Accuracy for Different Type of Optimization Methods and ResNet models

Optimization Method	Accuracy					
	ResNet18		ResNet50		ResNet101	
	Training	Testing	Training	Testing	Training	Testing
adam	99.1%	98.4%	99.8%	99.4%	99.3%	99.0%
sgdm	98.7%	98.8%	98.5%	98.1%	99.1%	98.9%
rmsprop	97.2%	96.8%	97.2%	96.8%	98.2%	98.7%

The dataset images were preprocessed before training the models. The training was conducted for 10 epochs with 200 iterations per epoch. Figure 6 illustrates the ResNet-50 network training performance using the Adam optimization algorithm. The top graph plots the accuracy percentage against iterations, which rapidly increases and levels off near 100%. The bottom graph plots the losses, which decrease drastically and continuously, leveling off at very low error percentages toward the end of training.

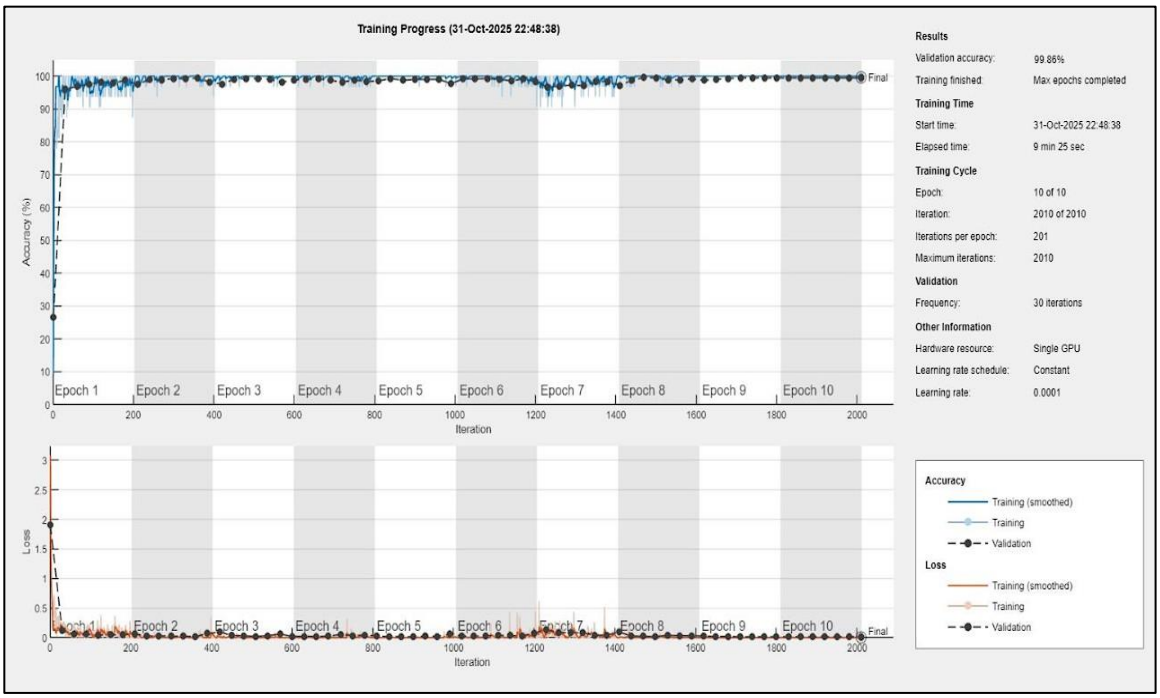
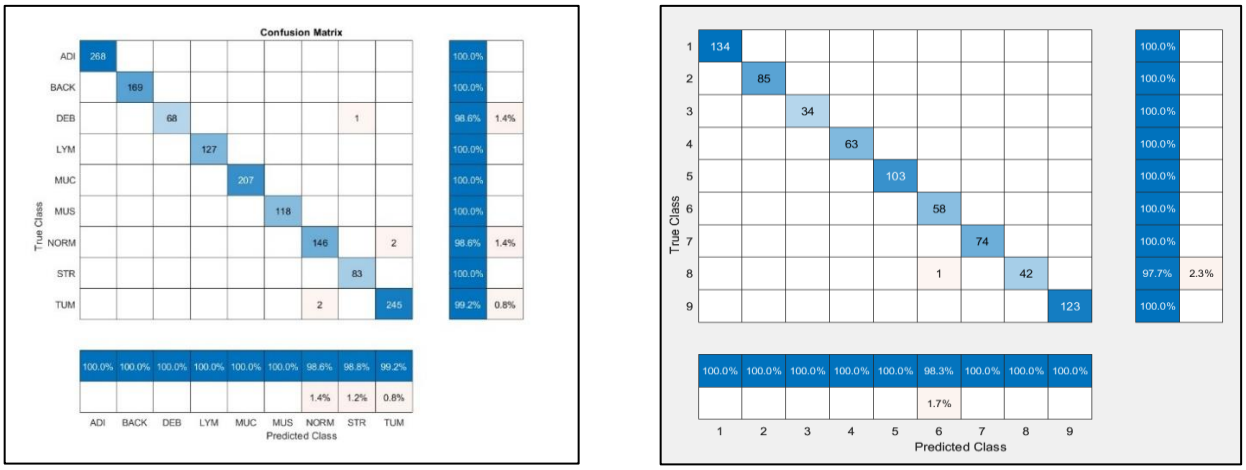


Figure 6: Epoch trends of the ResNet-50 network: Upper: epoch-accuracy plot; Bottom: epoch-loss plot. for NCT-CRC-HE-7K

Looking Figure 7,a: Under the **80/20 split**, the fine-tuned ResNet-50 achieves perfect classification for most classes; only minor errors are observed—**DEB 1.4%**, **MUS 1.4%**, **NORM 1.2%**, and **TUM 0.8%**—confirming strong class separability. While Figure 7,b: Under the **90/10 split**, the fine-tuned ResNet-50 shows near-perfect classification. Only **class 8** exhibits a single error (**2.3%** for that true class), and the predicted **class 6** shows a **1.7%** off-diagonal rate. All other classes reach **100%**.



(a) **(b)**

Figure 7: Confusion matrix of the fine-tuned ResNet-50 model for the NCT-CRC-HE-7K dataset a: train/test 80/20, b: train/test = 90/10

Table 3 reports per-class precision, recall, and F1-score on NCT-CRC-HE-7K. All classes achieve 100% across metrics except **MUS** (Precision = 95.1%, Recall = 98.3%, F1 = 96.7%) and **STR** (Recall = 97.7%, F1 = 98.8%), confirming near-perfect class separability.

Table 3: Classification Performance Metrics per Class (Precision, Recall, F1-Score) on NCT-CRC-HE-7K

Class	Precision (%)	Recall (%)	F1-Score (%)
ADI	100.0	100.0	100.0
BACK	100.0	100.0	100.0
DEM	100.0	100.0	100.0
LYM	100.0	100.0	100.0
MUC	100.0	100.0	100.0
MUS	95.1	98.3	96.7
NURM	100.0	100.0	100.0
STR	100.0	97.7	98.8
TUM	100.0	100.0	100.0

The multiclass ROC analysis (Figure 8) shows that all classes achieve very high AUC (AUC ≈ 1.0 for most classes and 0.986 for the lowest), confirming strong separability between colorectal tissue types.

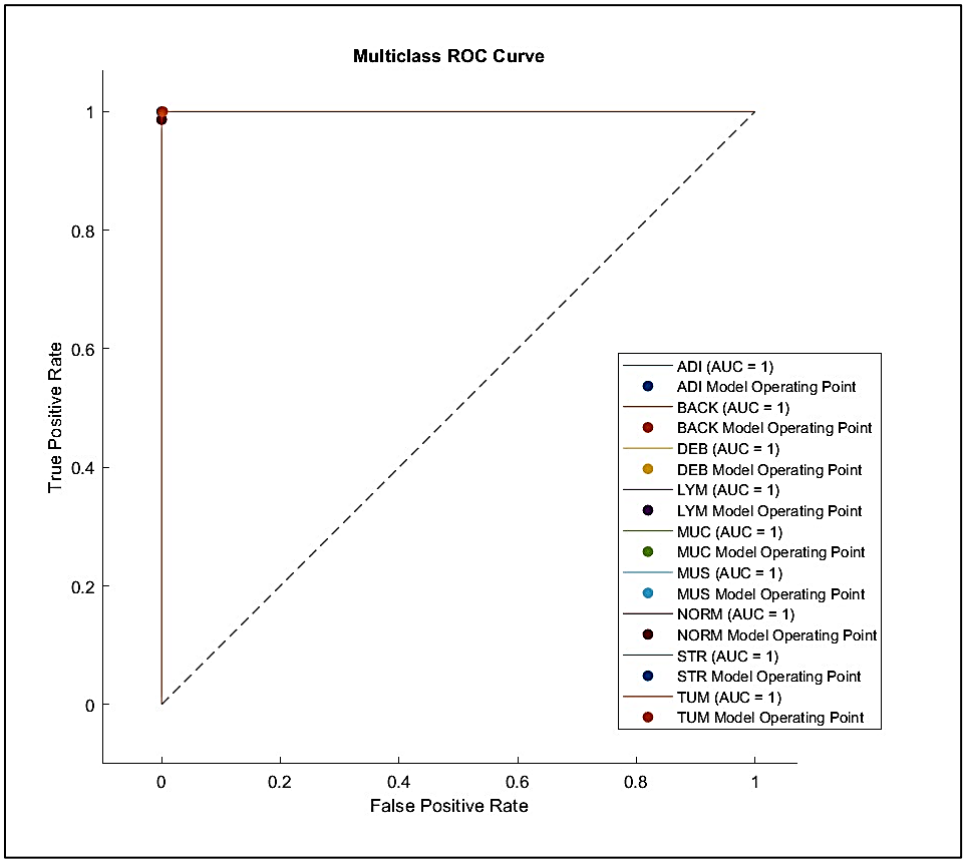


Figure 8: Multiclass ROC curves with per-class AUC on NCT-CRC-HE-7K

On NCT-CRC-HE-100K with a 90/10 split, the model attains 98.48% accuracy. Per-class metrics in Table 4 show uniformly high performance (overall Precision = 98.47%, Recall = 98.46%, F1 = 98.46%), with only small degradations for STR and TUM.

Table 4: Classification Performance Metrics per Class (Precision, Recall, F1-Score) on NCT-CRC-HE-100K

Class	Precision (%)	Recall (%)	F1-Score (%)
ADI	99.02	100.0	99.51
BACK	100.0	100.0	100.0
DEM	96.15	99.01	97.56
LYM	100.0	100.0	100.0
MUC	97.06	98.02	97.54
MUS	100.0	98.02	99.0
NURM	98.02	98.02	98.02
STR	97.98	96.04	97.0
TUM	98.00	97.03	97.51
Overall	98.47	98.46	98.46

Training converged rapidly on NCT-CRC-HE-100K: validation accuracy stabilized near 98–99% early and reached 98.36% at completion, with training/validation curves closely aligned and loss approaching zero—indicating stable learning without evident overfitting.

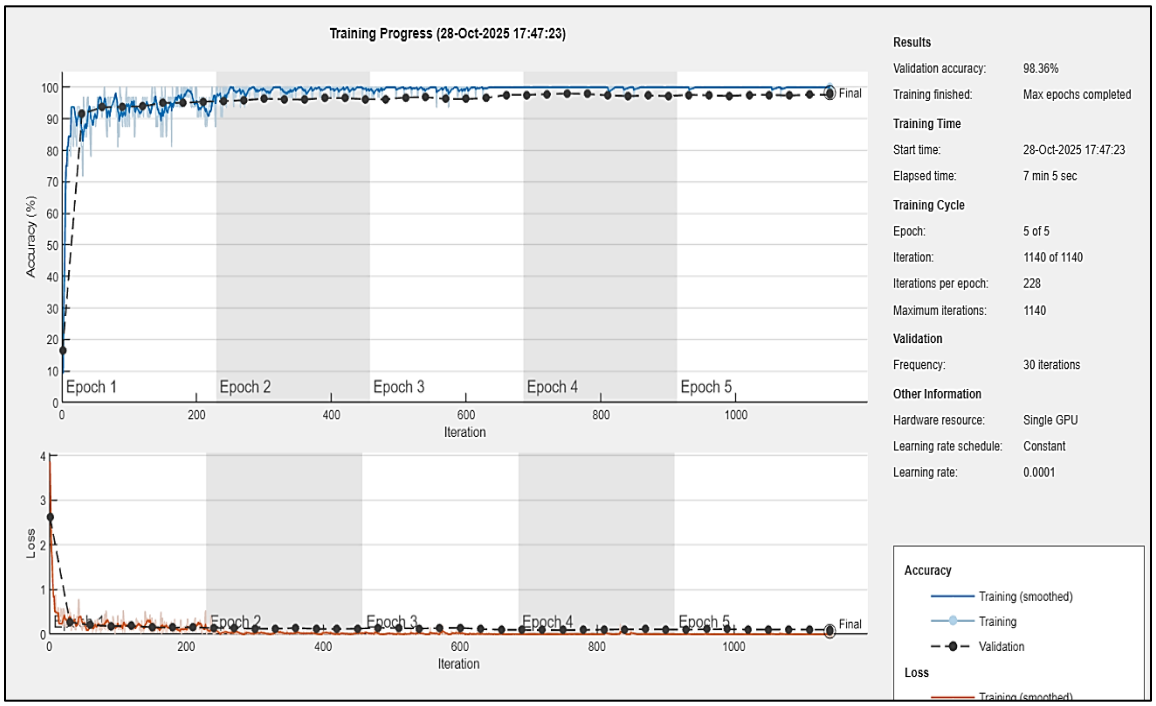


Figure 8: Epoch trends of the ResNet-50 network: Upper: epoch-accuracy plot; Bottom: epoch-loss plot. For NCT-CRC-HE-100K

On NCT-CRC-HE-100K, the multiclass ROC (Figure 9) shows near-perfect discrimination across classes (AUCs: ADI=1.000, BACK=1.000, LYM=1.000, NORM=0.9997, DEB=0.9996, MUS=0.9996, TUM=0.9988, STR=0.9982, MUC=0.9963), confirming robustness on the larger dataset.

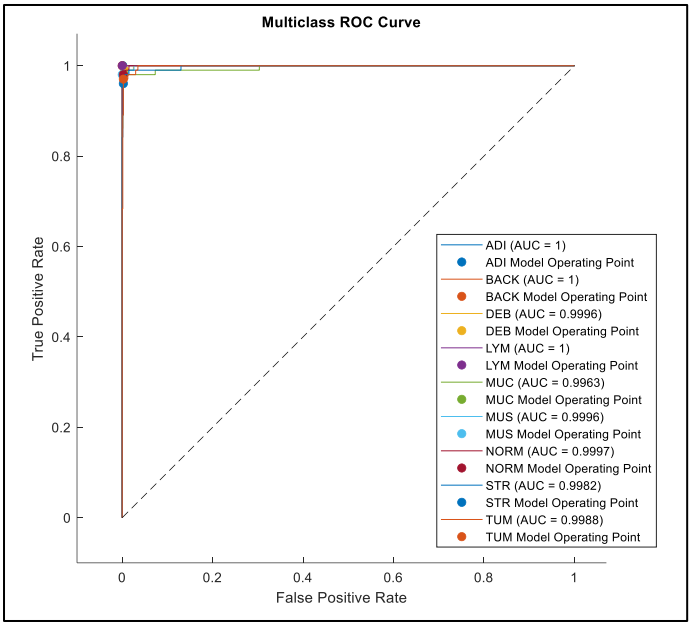


Figure 9: Multiclass ROC curves with per-class AUCs for NCT-CRC-HE-100K (90/10 split)

For NCT-CRC-HE-100K (90/10 split), the confusion matrix (Figure 10) shows strong diagonal dominance with small, sparse errors. Class-wise recall is: ADI 99.0%, BACK 100.0%, DEB 96.2%, LYM 100.0%, MUC 97.1%, MUS 100.0%, NORM 98.0%, STR 98.0%, TUM 98.0%. Column-wise precision remains high (ADI 100%, BACK 100%, DEB 99%, LYM 100%, MUC 98%, MUS 98%, NORM 98%, STR 96%, TUM 97%), confirming robust class separability with only minor confusions (≤4%).

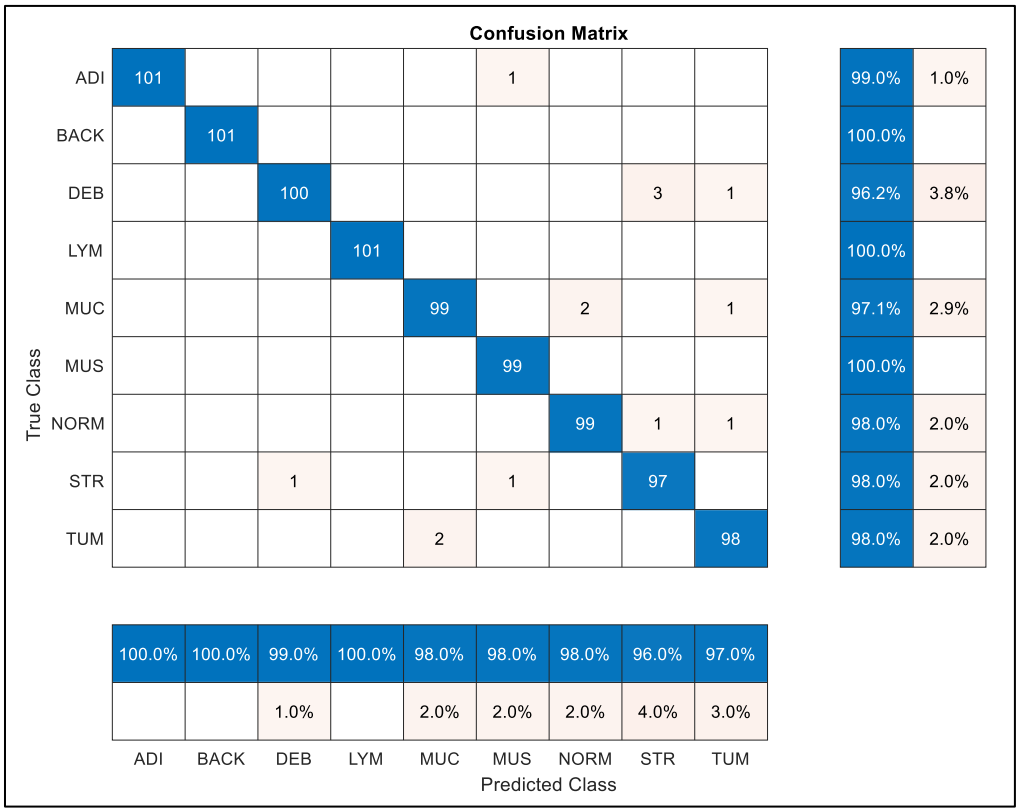


Figure 10: Confusion matrix of the fine-tuned ResNet-50 on NCT-CRC-HE-100K (train/test = 90/10)

Finally, the performance of our proposed method was compared with some state-of-the-art methods described in some previous research. Table 5 provides a comparative study of these methods based on their datasets, used techniques, and classification accuracies. Our proposed ResNet-50 model based on the NCT-CRC-HE-7K dataset and optimized by Adam achieved the highest accuracy of 99.86%, which was found to be superior to the other methods presented in the literature.

Table 5: Comparison between the proposed method and the current methods.

Authors	Dataset	Method	Result
Hamida et al. (2021)	NCT-CRC-HE-100K	ResNet	96.77%
Masud et al. (2021)	LC250000	Fully connected CNN	96.33%
Bukhari et al. (2020)	CRAG	ResNet-50	93.91%
Khan, et al. (2024)	LC25000	ResNet-101	99.80%
Ignatov, et al. (2024)	NCT-CRC-HE	EfficientNet-B0	97.7%
Uddin, et al. (2024)	NCT-CRC-HE-7K	CNN	99.5%
Proposed	NCT-CRC-HE-7K	ResNet-50	99.86%
	NCT-CRC-HE-100K	ResNet-50	98.36%

The superior accuracy of the proposed ResNet-50 model is mainly due to its residual learning framework, which enables effective training of deep networks and enhances feature extraction through skip connections. Its bottleneck design improves efficiency, while the Adam optimizer ensures faster convergence with adaptive learning rates. Additionally, preprocessing with median and sharpen filters improved image clarity, helping the model detect subtle patterns in colorectal cancer tissue. The high performance of the ResNet-50 + Adam configuration is not only numerical; it also has practical diagnostic value. The confusion matrix and class-level precision/recall show that the model correctly identifies colorectal cancer tissue patches with very few false negatives, which means that suspicious malignant regions are rarely missed. This is important for clinical triage, because in digital pathology workflows the first question is not “final diagnosis,” but “which regions should the pathologist look at first?” A model that reliably flags malignant-looking areas can reduce screening time in high-volume labs and help prioritize critical slides for manual review by the human expert. At the same time, the relatively low misclassification rate between normal and malignant tissue types suggests that such a system could be integrated as an assistive tool to focus attention rather than to replace expert judgment. We note that these results were consistent across folds, which indicates that the effect is stable and not due to a single favorable split.

5 Conclusion

This study compared deep learning methods for colorectal cancer detection using a comparison of different ResNet models (ResNet-18, ResNet-50, and ResNet-101) optimized with different optimization methods. The results indicated that an accuracy of 99.86% was achieved with the ResNet-50 model optimized using Adam on the dataset NCT-CRC-HE-7K and 98.36% on the dataset NCT-CRC-HE-100K and outperformed the other models and optimizers. The findings confirm the application of deep learning techniques to classify colorectal cancer, more specifically ResNet-50 under Adam optimization, to correctly discriminate with high precision and recall. This work complements existing research work in the field of digital pathology by describing an optimized model to classify colorectal cancer reliably and automatically. One limitation of this study is its reliance on a single dataset (NCT-CRC-HE-7K), which may affect the generalizability of the results. Future work will explore additional variations in datasets and design improvements to further enhance the model's efficiency and generalizability.

Acknowledgements

The authors have no acknowledgments to declare.

Authors contribution

Herman led the methodology and software, modified the network, trained/tested the models, and drafted the manuscript. Amira supervised the study and handled review and editing. Lozan performed data analysis and interpretation. Salwa collected the data and prepared tools/resources. All authors approved the final manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F.. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians, **71**(3), 209-249, doi: 10.3322/caac.21660 (2021)

[2] Mármol, I., Sánchez-de-Diego, C., Pradilla Dieste, A., Cerrada, E., & Rodriguez Yoldi, M. J. Colorectal carcinoma: a general overview and future perspectives in colorectal cancer. International journal of molecular sciences, **18**(1), 197, doi: 10.3390/ijms18010197 (2017).

[3] Granados-Romero, J. J., Valderrama-Treviño, A. I., Contreras-Flores, E. H., Barrera-Mera, B., Herrera Enríquez, M., Uriarte-Ruíz, K., ... & Arauz-Peña, G. Colorectal cancer: a review. Int J Res Med Sci, **5**(11), 4667, doi: 10.18203/2320-6012.ijrms20174914 (2017).

[4] Kather, J.N.; Krisam, J.; Charoentong, P.; Luedde, T.; Herpel, E.; Weis, C.-A.; Gaiser, T.; Marx, A.; Valous, N.A.; Ferber, D.; et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. PLoS Med. 2019, 16, e1002730, doi: 10.1371/journal.pmed.1002730

[5] Kather, J., Weis, CA., Bianconi, F. et al. Multi-class texture analysis in colorectal cancer histology. Sci Rep **6**, 27988, doi: 10.1038/srep27988 (2016).

[6] Komura, D., & Ishikawa, S. Machine learning methods for histopathological image analysis. Computational and structural biotechnology journal, **16**, 34-42, doi: 10.1016/j.csbj.2018.01.001 (2018).

[7] Alboaneen, D., Alqarni, R., Alqahtani, S., Alrashidi, M., Alhuda, R., Alyahyan, E., & Alshammari, T. Predicting colorectal cancer using machine and deep learning algorithms: Challenges and opportunities. Big Data and Cognitive Computing, **7**(2), 74, doi: 10.3390/bdcc7020074 (2023).

[8] Poston, G. J., Tait, D., O'connell, S., Bennett, A., & Berendse, S. Diagnosis and management of colorectal cancer: summary of NICE guidance. Bmj, 343, doi: 10.1136/bmj.d6751 (2011).

[9] Selvachandran SN, Hodder RJ, Ballal MS, Jones P, Cade D. Prediction of colorectal cancer by a patient consultation questionnaire and scoring system: a prospective study. Lancet. 2002;360, 278–283, doi: 10.1016/s0140-6736(02)09549-1

[10] Hippisley-Cox J, Coupland C. Identifying patients with suspected colorectal cancer in primary care: derivation and validation of an algorithm. Br J Gen Pract. 2012;62, e29–e37, doi: 10.3399/bjgp12x616346

[11] Chlorogiannis, D. D., Verras, G. I., Tzelepi, V., Chlorogiannis, A., Apostolos, A., Kotis, K., ... & Mulita, F. Tissue classification and diagnosis of colorectal cancer histopathology images using deep learning algorithms. Is the time ripe for clinical practice implementation?. Gastroenterology Review/Przegląd Gastroenterologiczny, **18**(4), 353-367, doi: 10.5114/pg.2023.130337 (2023).

[12] Jakhar, D., & Kaur, I. Artificial intelligence, machine learning and deep learning: definitions and differences. Clinical and experimental dermatology, **45**(1), 131-132, doi: 10.1111/ced.14029 (2020).

[13] Shen, D., Wu, G., & Suk, H. I. Deep learning in medical image analysis. Annual review of biomedical engineering, 19(1), 221-248, doi: 10.1146/annurev-bioeng-071516-044442 (2017).

[14] Pacal, I., Karaboga, D., Basturk, A., Akay, B., & Nalbantoglu, U. A comprehensive review of deep learning in colon cancer. Computers in Biology and Medicine, **126**, 104003, doi: 10.1016/j.compbiomed.2020.104003 (2020).

[15] Kumar, V. R. P., Arulselvi, M., & Sastry, K. B. S. Comparative assessment of colon cancer classification using diverse deep learning approaches. Journal of Data Science and Intelligent Systems, **1**(2), 128-135, doi: 10.47852/bonviewJDSIS32021193 (2023).

[16] Sirinukunwattana, K., Domingo, E., Richman, S. D., Redmond, K. L., Blake, A., Verrill, C., ... & Koelzer, V. H. Image-based consensus molecular subtype (imCMS) classification of colorectal cancer using deep learning. Gut, **70**(3), 544-554, doi: 10.1136/gutjnl-2019-319866 (2021).

[17] Bokhorst, J. M., Nagtegaal, I. D., Fraggetta, F., Vatrano, S., Mesker, W., Vieth, M., ... & Ciompi, F. Deep learning for multi-class semantic segmentation enables colorectal cancer detection and classification in digital pathology images. Scientific Reports, 13(1), 8398, doi: 10.1038/s41598-023-35491-z (2023).

[18] Tsai, M. J., & Tao, Y. H. Deep learning techniques for the classification of colorectal cancer tissue. Electronics, **10**(14), 1662, doi: 10.3390/electronics10141662 (2021).

[19] Jiang, X., Hoffmeister, M., Brenner, H., Muti, H. S., Yuan, T., Foersch, S., ... & Kather, J. N. End-to-end prognostication in colorectal cancer by deep learning: a retrospective, multicentre study. The Lancet Digital Health, **6**(1), e33-e43, doi: 10.1016/s2589-7500(23)00208-x (2024).

[20] Prezja, F., Äyrämö, S., Pölönen, I., Ojala, T., Lahtinen, S., Ruusuvuori, P., & Kuopio, T. Improved accuracy in colorectal cancer tissue decomposition through refinement of established deep learning solutions. Scientific Reports, **13**(1), 15879, doi: 10.1038/s41598-023-42357-x (2023).

[21] Masud, M., Sikder, N., Nahid, A. A., Bairagi, A. K., & AlZain, M. A. A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. Sensors, **21**(3), 748, doi: 10.3390/s21030748 (2021).
[22] Nemlander, E., Ewing, M., Abedi, E., Hasselström, J., Sjövall, A., Carlsson, A. C., & Rosenblad, A. A machine learning tool for identifying non-metastatic colorectal cancer in primary care. European Journal of Cancer, **182**, 100-106, doi: 10.1016/j.ejca.2023.01.011 (2023).
[23] Bostanci, E., Kocak, E., Unal, M., Guzel, M. S., Acici, K., & Asuroglu, T. Machine learning analysis of RNA-seq data for diagnostic and prognostic prediction of colon cancer. Sensors, **23**(6), 3080, doi: 10.3390/s23063080 (2023).
[24] Kather, J. N., Halama, N., & Marx, A. 100,000 histological images of human colorectal cancer and healthy tissue (v0.1) [Data set]. Zenodo, doi: 10.5281/zenodo.1214456 (2018).
[25] He, K., Zhang, X., Ren, S., & Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (770-778), doi: 10.1109/CVPR.2016.90 (2016).
[26] Khan, A. A., Arslan, M., Tanzil, A., Bhatty, R. A., Khalid, M. A. U., & Khan, A. H. Classification of colon cancer using deep learning techniques on histopathological images. Migration Letters, **21**(S11), 449-463. https://migrationletters.com/index.php/ml/article/view/10719/7163, (2024).
[27] Ignatov, A., & Malivenko, G. NCT-CRC-HE: Not All Histopathological Datasets Are Equally Useful. arXiv preprint arXiv:2409.11546, doi: 10.48550/arXiv.2409.11546 (2024).
[28] Uddin, A. H., Chen, Y. L., Akter, M. R., Ku, C. S., Yang, J., & Por, L. Y. Colon and lung cancer classification from multi-modal images using resilient and efficient neural network architectures. Heliyon, **10**(9), doi: 10.1016/j.heliyon.2024.e30625 (2024).