

Histopathological cancer detection based on deep learning and stain images

Dina M. Ibrahim^{1,2}, Mohammad Ali A. Hammoudeh¹, Tahani M. Allam²

¹Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia

²Department of Computers and Control Engineering, Faculty of Engineering, Tanta University, Tanta, Egypt

Article Info

Article history:

Received Feb 3, 2024

Revised Apr 15, 2024

Accepted May 12, 2024

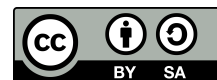
Keywords:

Colorectal cancer
Convolutional neural network
Deep learning
Detection method
Machine learning

ABSTRACT

Colorectal cancer (CRC)-a malignant growth in the colon or rectum- is the second largest cause of cancer deaths worldwide. Early detection may increase therapy choices. Deep learning can improve early medical detection to reduce the risk of unintentional death from an incorrect clinical diagnosis. Histopathological examination of colon cancer is essential in medical research. This paper proposes a deep learning-based colon cancer detection method using stain-normalized images. We use deep learning methods to improve detection accuracy and efficiency. Our solution normalizes image stain variations and uses deep learning models for reliable classification. This research improves colon cancer histopathology analysis, which may enhance diagnosis. Our paper uses DenseNet-121, VGG-16, GoogLeNet, ResNet-50, and ResNet-18 deep learning models. We also analyze how stain normalization (SN) improves our model on histopathology images. The ResNet-50 model with SN yields the highest values (9.94%) compared to the other four models and the nine models from previous studies.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Dina M. Ibrahim

Department of Information Technology, College of Computer, Qassim University

Buraydah, Saudi Arabia

Email: d.hussein@qu.edu.sa

1. INTRODUCTION

Uncontrollable abnormal cell development is the root cause of cancer. It can start in any organ or tissue of the body and metastasize (spread) throughout. Cancer is the second-leading cause of death around the globe [1]. The incidence and mortality rates of cancer are growing rapidly and globally. The increasing rates are reflected in increasing lifespans and population growth [2].

Colorectal cancer (CRC) is the development of cancer in the colon or rectum region [3], and it ranks second in terms of mortality [2]. A CRC is a type of tumor; however, not all tumors are cancerous. There are benign tumors where abnormal growth is not deadly and does not spread as much as a malignant tumor, which is more spreadable, destructive, and, in most cases, deadly. Most cases start as benign tumors, and over time, some of those tumors become malignant [3]. Colon cancer is often referred to as CRC because of its similarity with rectum cancer. In the case of colon cancer, adenocarcinoma is the most common type of cancer, which starts in the glands of the organ. Adenocarcinoma can also start in the breasts, lungs, and pancreas. Certain lifestyle factors might increase the risk of colon cancer, for example, obesity, consuming high amounts of red meat, low physical activity, smoking, and alcohol consumption. Other risk factors are not lifestyle-related, like being male or older, and others are hereditary, like a family history of some inherited genetic disorders.

While there are no conclusive symptoms of CRC, there are some warning signs, such as blood in the stool, severe constipation, weight loss, and appetite loss. Some diagnostic techniques include X-rays, magnetic resonance imaging (MRI), and computed tomography (CT) scans, which are used to look for suspicious areas and see how far cancer has spread. A colonoscopy test is done by passing a thin, lighted tube with a small camera through the colon to screen for any abnormalities. A biopsy is done when other tests suggest the presence of suspicious tissues by examining a small amount of the suspected tissue under the microscope. Treatments depend on the stage of cancer, but treatment options include surgery to remove the polyp during colonoscopy, radiation therapy, or chemotherapy to destroy or shrink cancer cells [4].

Machine learning (ML) has been applied increasingly in the medical field to aid with the detection, classification, and survival rate prediction of diseases, cancer in particular, whether in decision-making training or automated diagnosis, for the ultimate purpose of preserving life. Deep learning and convolutional neural networks (CNN) are being progressively used for optical pathology image analysis, improving the performance of colon cancer diagnosis. The subfield of ML known as deep learning is employed in medical image processing, computer vision, audio analysis, and natural language processing. One of the most popular AI tools. It involves arranging linear and non-linear processing units in deep architecture to automatically extract valuable characteristics [5]. Deep learning advances the classical neural network methodology. In the past six years, computer power and new data sets have revived it. Deep learning analyzes complex, non-linear data. It has been used successfully in medicine to diagnose diseases.

Deep CNNs have excelled in automated vision and image processing competitions. CNN is promising for categorization and segmentation, recognizing objects, processing videos, natural language interpreting, and recognition of speech. The stacking of more than one linear and non-linear processing unit in a layering fashion provides the ability to understand complicated representations at different abstraction levels, giving deep architectures an advantage over shallow ones. The arrival of deep learning has transformed the image analysis field. Using neural networks, which can recognize objects based on the features they have learned from training the data, it is possible to solve complex visual tasks.

Bad clinical choices due to a lack of experience may lead to an unexpected loss of life. Thus, the intelligent decision support system (DSS) is not meant to substitute a doctor but to assist him with decision-making training for specific diseases [5]. In health services such as cancer screening and diagnosis, disease control, and treatment approaches, ML practices are used. Deep learning is at the forefront of ever-increasing medical data and transforming this data into useful knowledge. The efficiency of deep learning is truly exceptional in diagnosing colon cancer. Success in subjects such as pathology image analysis or colonoscopy image analysis has a significant role in the diagnosis of colon cancer. Current clinical dataset-based COVID-19 mortality risk prediction studies use ML algorithms with little or no deep learning. Additionally, most such investigations use a tiny clinical dataset. By using the medical dataset in both CSV and converted picture formats, we can improve deep learning model performance. To this end, our study introduces three deep-learning predictive models for COVID-19 mortality risk prediction. The main contributions of this study are:

- Developing five different models of deep learning with different architectures to detect histopathological Colon Cancer.
- Using the stain-normalized images instead of the original images and examine the effectiveness of the stain-normalization.
- Conducting a comparison between the five models and determining which is the best model.
- Comparing the results of our proposed work with previous studies that used deep learning CNN models.

The remainder of the paper is organized as follows: section 2. illustrates a literature review on colon cancer histopathology. Materials and methods are described in section 3. which presents the steps of our histopathological colon cancer detection model. In section 4. explains the results of our proposed models with more discussions. Finally, section 5. draws the concluding remarks.

2. LITERATURE REVIEW ON COLON CANCER HISTOPATHOLOGY

All through our research, we floundered upon many topics in the field of histopathological cancer detection using deep learning. In this section, we review the most eminent research we have found on our topic, histopathological colon cancer detection using deep learning. Ke *et al.* [6], proposed a high-throughput analysis system for whole slide images (WSI) to localize tumor regions. The architecture is constructed with Monte

Carlo (MC) distribution approximation, a conditional random field (CRF)-based tissue classifier, and DenseNet as the backbone. The system achieved 97.34% classification accuracy with DenseNet and an area under the curve (AUC) of 87.3%. The study used three colorectal datasets provided by the cancer genome atlas (TCGA), consisting of WSI. The datasets consisted of three subsets i) Stomach adenocarcinoma (TCGA-STAD) with 432 samples; ii) Colon adenocarcinoma (TCGA-COAD) with 460 samples; iii) Rectum adenocarcinoma (TCGA-READ) with 171 samples.

The study presented by Awan *et al.* [7] suggested employing a unique two-stage CNN-based framework for the classification of picture patches, thereby automating the examination of pre-cancerous tissue abnormalities. The framework will analyze the input to identify the glandular structures, and then, the processed image will be inputted into a classification model to forecast any anomalies. The researchers reduced the correlation between the CRC pictures by removing the non-glandular regions of the photos. The gland segmentation on a large dataset of WSI images was accomplished by training a CNN model. Subsequently, a CNN was employed for classification in order to acquire knowledge about the characteristics of the glandular structures. The dataset was obtained by capturing images through the use of a commercially available WSI scanner. The dataset consists of pictures from three subsets: i) normal, ii) carcinoma (CA), and iii) (TA.LG). The evaluation of the classification methods is categorized into three groups of images: i) unprocessed images (Iraw), ii) images with ground truth segmentation (IGTS), and iii) images with network segmentation (INS). The study obtained optimal categorization outcomes while employing GoogleNet. The results per setting were: i) For Iraw, the F1-score obtained is 82.9%, and the accuracy is 91.5%; ii) For IGTS, the F1-score obtained is 87.1%, and the accuracy is 95.7%; iii) For INS, the F1-score obtained is 84.9%, and the accuracy is 92.5%.

Kosaraju *et al.* [8] introduced Deep-Hipo, a multi-task-based deep learning model, enabling accurate picture interpretation by simultaneously analyzing multi-scale patches. In a WSI, Deep-Hipo extracts two patches of the same window at high and low magnification levels, capturing complex morphological features in both fields. The tissue samples were stained with H&E utilizing 94 samples. Deep-Hipo calculates cancer likelihood on each pixel, slides multi-scale windows over a histopathology picture, and visualizes probability scores to locate cancerous spots. Deep-Hipo creates multi-scale receptive fields by concurrently obtaining two patch photos (at high and low magnification) from two CNN tracks and concatenating the last hidden layers to calculate cancer risk. High-magnification patch images are captured on one track (20×). A track with a low magnification (5×) captures two patch images of the same pixel size from the same central position in a WSI. Each track has 21 layers. The last Inception modules of both songs are combined. The completely linked layer receives the flattened layer. The output layer has one binary classification node to estimate posterior probability, indicating if a patch is malignant. Deep-hipo was performed on well-differentiated, moderately differentiated, and poorly differentiated adenocarcinomas, yielding 93.7% accuracy, 93.4% F1-score, 92.9% precision, and 93.9% recall.

Dif and Elberichi [9] suggested a dynamic ensembling method for choosing the best modules. They tested it on eight CNN architectures: Inception-ResNet, Xception, MobileNet, VGG, Inception, ResNet, and DenseNet. They used a public dataset that consists of 5000 histopathological images of 150×150 that have been extracted from 10 WSI to classify CRC into 27 types of tissue classes. The study used an ensembling approach to achieve three main goals. They started by transferring all the conclusion layers from the pre-trained models on the image in a dataset to the new model to prevent overfitting and reduce computational complexity. The second step was to feed the training and validation sets to the model generator. This step creates a collection of 100 models resulting from dynamic selection; they proposed two strategies to improve performance and diversity for the second step. The first one was to use 90% of the training set to ensure diversification, and the second was to roll out weak learners keeping models that achieved high validation accuracy; Step two involves narrowing down the created set of models to the most effective Asian company. Last but not least, we evaluated how well the chosen models worked together using the practical swarm optimization (PSO) technique for dynamic ensembling. Thanks to the Resnet121 framework and voting mechanism, the method has attained an accuracy rating of 94.52%.

Yoon *et al.* [10] the author's development of five different modifications on the VGG architecture, one of the CNN networks, they studied the effect that the depth of a CNN has on its performance, specifically the detection accuracy. They use a private dataset composed of 28 normal and 29 tumor images sized 12,762x13,939 px that were cropped to 256x256 px giving a complete dataset of 10,280 smaller images consisting of 6806 normal and 3470 tumor images. They created VGG models from A to E and divided that implementation into two steps; for the first step, they used 400 images divided equally between normal and tumor images; these are

the following accuracies for each VGG composition A–E, 82.5%, 87.5%, 87.5%, 91.4%, and 94.3%, in that order. The second step used the VGG with the highest accuracy on the complete 10,280 image dataset, which was VGG E with a 94.3% score on accuracy.

Sarwinda *et al.* [11] compared the performance of two ResNet architectures, ResNet-18 and ResNet-50, for detecting CRC. The models have been trained to classify CRC gland images into benign and malignant categories. The authors compared using three different training data to testing data ratios 60%:40%, 75%:25%, and 80%:20%. For image preprocessing, the authors employed a grayscale image formation process and a contrast-limited adaptive histogram equalization (CLAHE) to improve the contrast in the images. The experiment results showed that ResNet-50 was more effective than ResNet-18, achieving an accuracy of 88%, sensitivity of 93%, and specificity of 83%. The highest accuracy was obtained using the ratio of 80% of training data and 20% of testing data. In Table 1, we summarized our findings and highlighted similar aspects for comparison.

Table 1. Literature review summary of histology colon cancer studies

Reference	Architecture	Dataset	Metrics
[6]	DenseNet	Public	Accuracy = 97.34% AUC = 87.3%
[7]	GoogLeNet	Private	Accuracy = 91.5% F1-score = 82.9%
[8]	Dual-Channel crafted CNN	Private	Accuracy = 93.7% F1-score = 93.4% Precision = 92.9% Recall = 93.9%
[9]	ResNet-121	Public	Accuracy = 94%
[10]	Crafted VGG	Private	Accuracy = 93.48% Sensitivity = 95.10% Specificity = 92.76%
[11]	ResNet-50 & ResNet-18	Public	Accuracy = 88% Sensitivity = 93% Specificity: 83%
[12]	RCCNet (CNN based)	Public	Accuracy = 80.61% Average F1 = 78.87%
[13]	SVM + GLCM	Public	Precision = 96.67% recall = 83.33% F-measures = 89.51%
[14]	CNN + SampEn	Public	AUC = 98.3%
[15]	DenseNet-121 with K-nearest neighbor (KNN) classifier	Public	Accuracy = 98.53% Sensitivity = 98.63% Specificity = 98.43%

Basha *et al.* [12], the authors suggested RCCNet, a CNN architecture for the categorization of routine colon carcinoma nuclei based on histology. The model consists of 1, 512, and 868 learnable parameters, unlike other CNN models that consist of significantly more. The authors compared the proposed RCCNet model with five CNN models, which are GoogLeNet, AlexNet, CIFAR-VGG, WRN, SoftmaxCNN, and SoftmaxCNN IN27 in terms of accuracy and average F1-score. The proposed model outperformed the other four CNN models by obtaining an accuracy of 80.61% and an average F1-score of 78.87%.

Jiao *et al.* [13] aimed to enhance the efficiency of colon cancer detection, where support vector machine (SVM) feature extraction was key in the research. They also used the gray-level co-occurrence matrix (GLCM) to extract 18 features for the classification and to obtain the classification results, they used three-fold validation. They used SVM for image classification and mean and variance features with GLCM features for future extraction. They used a dataset of 60 whole-slide histopathologic images including 30 normal and 30 abnormal images, that were provided privately by the hospital and labeled by doctors. Using the cross-validation method to improve the performance, using all of the 18 features, it's a cheese I precision of 96.67% a recall of 83.33% and finally, an F-measures of 89.51% improving the early detection of colon cancer.

A new method for assessing and classifying H&E histological pictures of CRC was created by the authors in [14]. They used SampEn MF, a combination of multivariate and fuzzy sample entropy. Images from sixteen WSI H&E were used to conduct the study. They used 67 benign and 84 malignant cases, adding up to 151 RGB images with dimensions of 775x522 pixels. To make a model, they started by splitting the images into areas with random windows. Then, they used a multiscale approach to look for pattern similarity by looking at windows of different sizes and variations. This was done by treating each pixel as an n-dimensional picture and analyzing it from the Minkowski distance. Next, they used a Gaussian function to check the relevance between windows, which made a method that can figure out pixel similarity between windows of different sizes. Finally, the results were given to a classifier with logistic scoring the highest score, which was an AUC of 98.3%.

Sarwinda *et al.* [15] proposed a feature extraction method using CNN models, ResNet-50 and DenseNet-121. To classify colon cancer histopathology images into benign and adenocarcinoma, the authors tested popular classifiers such as SVM, Random Forest, KNN, and XGBoost. The study evaluated the models on two ratios of training data to testing data (75%:25% and 85%:15%). The data used in the study was obtained from the Kaggle Lung and Colon Cancer Histopathological Image Dataset, which contains 10000 colon cancer images split into two classes: adenocarcinoma and benign. The experiment results showed that DenseNet-121 achieved higher accuracy, sensitivity, and specificity than the ResNet-50 architecture for all classifiers. DenseNet-121 achieved about 98.53% accuracy and 98.63% sensitivity with the KNN classifier. Literature review summary of histology colon cancer datasets are presented in Table 2.

Table 2. Literature review summary of histology colon cancer datasets

Reference	Dataset name	No. of images	Size (WxH)	No. of classes
[16]	Colon and Lung Cancer Histopathological Image Dataset	10,000	768x768	2
[17]	The cancer genome atlas (TCGA)	1,063	512x512	3
[18]	CRC dataset	5,000	150x150	7
[19]	The Warwick-QU Dataset	165	567x340	2
[20]	CRCHistoPhenotypes	100	500x500	4
[21]	The glas challenge contest dataset	151	775x522	2

3. MATERIALS AND METHODS

This section presents our proposed histopathological colon cancer detection model illustrated in Figure 1; our model is divided into four different steps. In this section, we explore those steps in detail, illustrating the expected results and the reasoning behind our decisions. We have selected the Colon and Lung Cancer Histopathological Image Dataset (LC25000) for our study [16]. The selection was made based on the dataset's magnitude and the diversity of its classes. The dataset consisted of two distinct classes, each including 5,000 photos. The first class, "colon," comprised 5,000 images of colon cancer, while the second class, "colon_n," consisted of 5,000 images of benign colonic tissues. The photos are resized to a square shape with dimensions of 768x768 pixels relative to their initial dimensions of 1,024x768 pixels. A subset of the dataset is displayed in Figure 2.

The colon and lung cancer histopathological images dataset is available publicly on Kaggle. This dataset contains 25,000 already augmented images split into two folders, the lung image set and the colon image set, as shown in Figure 1 samples of the dataset with labels: 0 for adenocarcinoma and 1 for benign. Our target of interest is the colon image set, This includes 10,000 JPEG photos with a resolution of 768 by 768 pixels. The colon image set is divided into two categories: adenocarcinoma and benign, separated by folders. Our plan is to create separate folders for training, testing, and validation:

- Throughout training and fitting, the model will learn the labeled data from this collection, which is called the training set.
- Set for testing: this set will be utilized in training and fitting the model, with the goal of improving the model's performance.
- Validation set: we will assess, validate, and test our final, trained model using this collection.

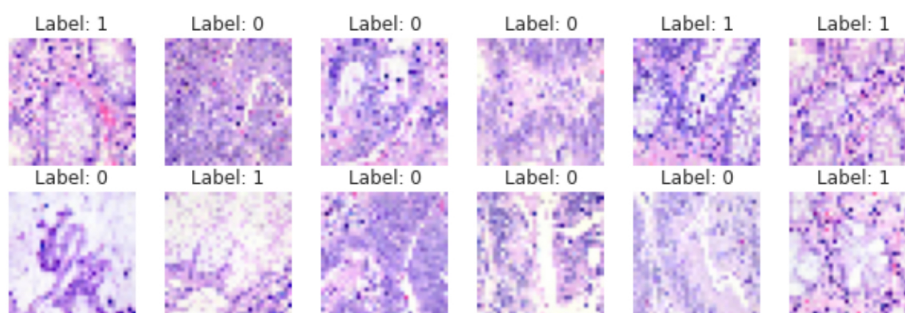


Figure 1. Samples of the dataset with labels: 0 for adenocarcinoma and 1 for benign

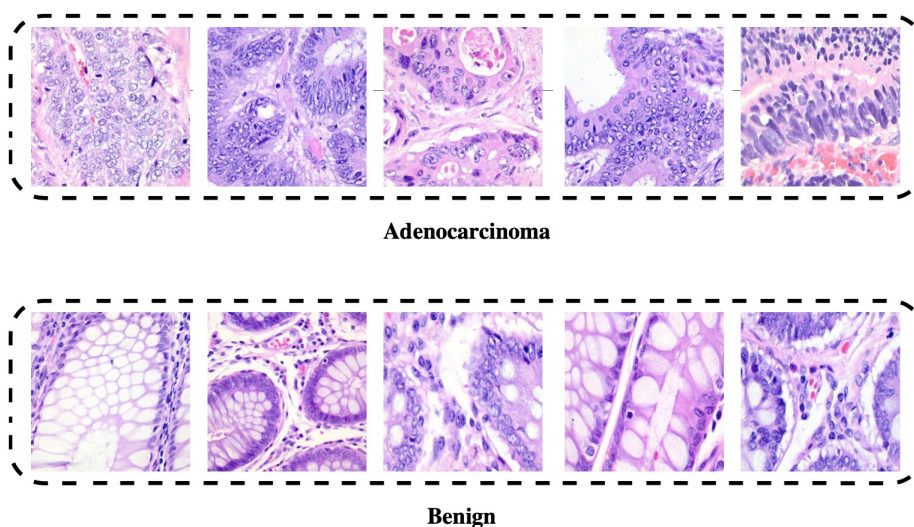


Figure 2. Samples of images found in the dataset: 1st row colon_aca (colon adenocarcinomas), 2nd row colon_n (benign colonic tissues)

3.1. Image preprocessing

In our model, we adopted different image preprocessing techniques, such as SN and data augmentation. Those steps were adopted to prepare the raw data and make it suitable for training the model. The steps are as follows:

- Stain normalization (SN) is a necessary step that we plan to incorporate in our model. Image classification needs an SN step to perform more efficiently, SN deals with stain variations and inconsistencies in hematoxylin and eosin (H&E) stained histopathology images, which is a major issue that can affect the results of the automatic image analysis algorithms [22]. The stain variations can be caused by using different staining chemicals, different color concentrations, or different scanners [23]. Ciompi [22] states that SN algorithms are concerned with converting an image I into a target image \hat{I} by a predefined template image Θ and comparing the way an input image looks to the way a template image looks using a mapping function f , by the formula $\hat{I} = f(I, \Theta)$. Figure 3 shows the difference between an image before and after being stain normalized [24], [25].
- Image augmentation techniques are for enhancing the size and quality of the training dataset [26]. The image augmentation techniques, including, but not limited to: geometric transformations, image flipping, cropping, rotation, zooming, translation, and noise injection, are illustrated in Figure 4. Image augmentation is a useful and important step that we use in this paper to produce new and different histopathology images from the already existing dataset.

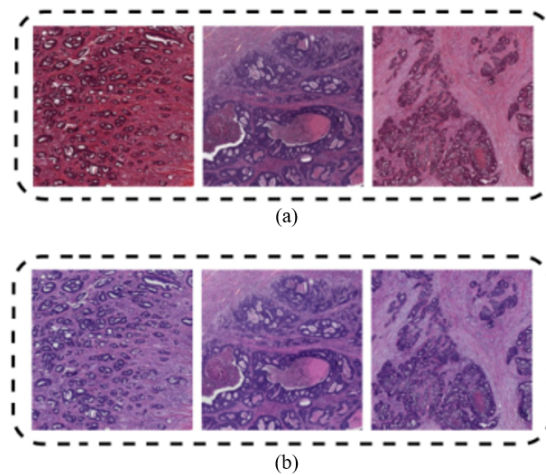


Figure 3. Samples of histopathology images (a) before SN and (b) after SN

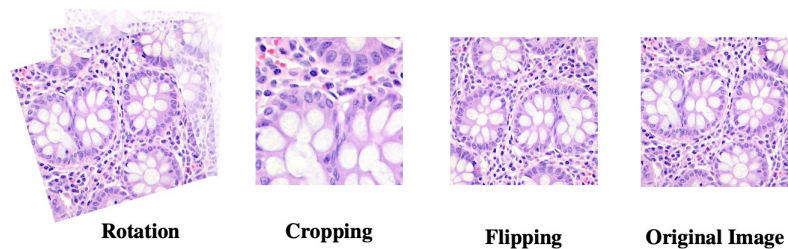


Figure 4. Image augmentation example

3.2. Feature extraction by deep learning models

A system can autonomously identify the representations needed for feature identification, prediction, or classification from the preprocessed dataset through feature learning, a collection of methods used for machine learning ML [27]. This enables a computer to understand the characteristics and use them to reproduce a particular task, such as prediction or classification. In deep learning, by developing a complete CNN, feature learning can be accomplished [28]. In this section, we build and compare five different CNN models and their performance. We chose to conduct our feature extraction using five different CNN networks: DenseNet, VGG, GoogLeNet, and ResNet, with a histopathology image dataset. Our selection of them was dependent on previously achieved results by previous research in the field of histopathological colon cancer detection

3.2.1. DenseNet-121 deep learning model

In comparison with AlexNet, GoogLeNet, and ResNet50, dense convolutional network trains with fewer parameters and better accuracy. It employs fundamental connection design for maximizing the transfer of data in either forward and the reverse pass computation to vanishing input and information gradients in the deep network. Each layer receives input maps of features from previous layers and outputs to subsequent layers. Features are reused across the network for compact learning. Layers of DenseNet are tightly connected. Transition layers within these blocks perform convolution and pooling. All layers are dense blocks with batch normalization, ReLu nonlinear activation function, and 3x3 convolution size [29].

DenseNet-121 finds 121 by multiplying 5 by $(6+12+24+16)*2$: 121. This means that there are 5 layers for convolution or pooling, 3 layers for transition (6,12,24), and 1 layer for classification (16). The result is then multiplied by 2 dense blocks (1x1 conv and 3x3 conv). Figure 5 represents the DenseNet-121 architecture in detail.

Layers	Output Size	DenseNet-121
Pooling	112 x 112	7 x 7 conv, stride 2
Convolution	56 x 56	3 x 3 max pool, stride 2
Dense Block (1)	56 x 56	$\begin{bmatrix} 1 \times 1 & conv \\ 3 \times 3 & conv \end{bmatrix} \times 6$
Transition Layer (1)	56 x 56	1 x 1 conv
	28 x 28	2 x 2 max pool, stride 2
Dense Block (2)	28 x 28	$\begin{bmatrix} 1 \times 1 & conv \\ 3 \times 3 & conv \end{bmatrix} \times 12$
Transition Layer (2)	28 x 28	1 x 1 conv
	14 x 14	2 x 2 max pool, stride 2
Dense Block (3)	14 x 14	$\begin{bmatrix} 1 \times 1 & conv \\ 3 \times 3 & conv \end{bmatrix} \times 24$
Transition Layer (3)	14 x 14	1 x 1 conv
	7 x 7	2 x 2 max pool, stride 2
Dense Block (4)	7 x 7	$\begin{bmatrix} 1 \times 1 & conv \\ 3 \times 3 & conv \end{bmatrix} \times 16$
Classification Layer	1 x 1	7 x 7 global average pool
		1000D fully-connected, softmax

Figure 5. DenseNet-121 architecture summary

3.2.2. GoogLeNet deep learning model

LeNet is one of the earliest CNN architectures that has been introduced by LeCun *et al.* [30], and it is considered the history of deep CNN. LeNet-5 is famous for its exemplary performance in recognizing handwritten digits and classifying them successfully. The classic architecture became an inspiration for future spatial exploitation-based CNNs like AlexNet and VGG [31]. GoogLeNet is a CNN, and it is a variant of the Inception Network. GoogLeNet is 22 layers deep; part of these layers are nine inception modules. Figure 6 represents the GoogLeNet architecture.

type	patch size	stride	output size	depth	#1x1	#3x3 reduce	#3x3	#5x5 reduce	#5x5	pool proj	params	ops
convolution	7 x 7	2	112 x 112 x 64	1							2.7K	34M
max pool	3 x 3	2	56 x 56 x 64	0								
convolution	3 x 3	1	56 x 56 x 192	2	64	192					112K	360M
max pool	3 x 3	2	28 x 28 x 192	0								
inception (3a)			28 x 28 x 256	2	64	96	128	16	32	32	159K	128M
inception (3b)			28 x 28 x 480	2	128	128	192	32	96	64	380K	304M
max pool	3 x 3	2	14 x 14 x 480	0								
inception (4a)			14 x 14 x 512	2	192	96	208	16	48	64	364K	73M
inception (4b)			14 x 14 x 512	2	160	112	224	24	64	64	437K	88M
inception (4c)			14 x 14 x 512	2	128	128	256	24	64	64	463K	100M
inception (4d)			14 x 14 x 528	2	112	144	288	32	64	64	580K	119M
inception (4e)			14 x 14 x 832	2	256	160	320	32	128	128	840K	170M
max pool	3 x 3	2	7 x 7 x 832	0								
inception (5a)			7 x 7 x 832	2	256	160	320	32	128	128	1072K	54M
inception (5b)			7 x 7 x 1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7 x 7	1	1 x 1 x 1024	0								
dropout (40%)			1 x 1 x 1024	0								
linear			1 x 1 x 1000	1							1000K	1M
softmax			1 x 1 x 1000	0								

Figure 6. GooLeNet architecture summary

3.2.3. VGG-16 deep learning model

VGG is a spatial exploitation-based CNN proposed by Karen and Andrew [32], the architecture showed good image classification results. VGG is famous for its homogenous topology and simplicity. The model uses many 3x3 and 1x1 tiny filters with a 1 pixel convolution stride. Max pooling layers operate on a 2x2 pixel window with stride 2. Not all convolutional layers are followed. After a convolutional stack, three completely connected layers. Filters increase with model depth. Many VGG-Net versions have been built and assessed, named for layer count. VGG-16 has 16 learned layers and VGG-19 has 19 learned layers [33]. The VGG-16 model is a spatial exploitation based on CNN and has a total of 16 learned layers. The model is shown in detail in the built-in Figure 7.


```

Model: "model_1"
├─ Layer (type)      Output Shape      Param #
=====
input_2 (InputLayer)  [(None, 200, 200, 3)]  0
block1_conv1 (Conv2D)  (None, 200, 200, 64)   1792
block1_conv2 (Conv2D)  (None, 200, 200, 64)   36928
block1_pool (MaxPooling2D) (None, 100, 100, 64)   0
block2_conv1 (Conv2D)  (None, 100, 100, 128)  73856
block2_conv2 (Conv2D)  (None, 100, 100, 128) 147584
block2_pool (MaxPooling2D) (None, 50, 50, 128)   0
block3_conv1 (Conv2D)  (None, 50, 50, 256)   295168
block3_conv2 (Conv2D)  (None, 50, 50, 256)   590080
block3_conv3 (Conv2D)  (None, 50, 50, 256)   590080
block3_pool (MaxPooling2D) (None, 25, 25, 256)   0
block4_conv1 (Conv2D)  (None, 25, 25, 512)   1180160
block4_conv2 (Conv2D)  (None, 25, 25, 512)   2359808
block4_conv3 (Conv2D)  (None, 25, 25, 512)   2359808
block4_pool (MaxPooling2D) (None, 12, 12, 512)   0
block5_conv1 (Conv2D)  (None, 12, 12, 512)   2359808
block5_conv2 (Conv2D)  (None, 12, 12, 512)   2359808
block5_conv3 (Conv2D)  (None, 12, 12, 512)   2359808
block5_pool (MaxPooling2D) (None, 6, 6, 512)     0
=====
Total params: 14,714,688
Trainable params: 0
Non-trainable params: 14,714,688

```

Figure 7. VGG-16 architecture summary

3.2.4. ResNet-18 deep learning model

ResNet is a novel architecture introduced by He *et al.* [34]. The architecture uses a series of residual blocks and shortcuts, or “skip connections,” that skip from a stack of convolutional layers and directly connect to the outputs. In a residual block, each layer is fed to the next layer and simultaneously connects directly to the layer 2 or 3 hops away. ResNet has many versions depending on the number of layers, and each residual block is either two layers deep or three layers deep. Smaller networks, for instance, ResNet18 and ResNet34, use a two-layer residual block, while larger networks, for instance, ResNet50, ResNet101, and ResNet152, use a three-layer residual block. Residual networks are either one of two types [34], either two layers-deep or three layers-deep residual blocks. ResNet-18 falls into the first category, and the model is shown in detail and built-in in Figure 8.

Layers	Output Size	ResNet-18
conv1	112 x 112	7 x 7, 64, stride 2
conv2_x	56 x 56	3 x 3 max pool, stride 2
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
conv3_x	28 x 28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
conv4_x	14 x 14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
conv5_x	7 x 7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
	1 x 1	average pool, 100-d fully-connected, softmax
FLOPs		1.8x10 ⁹

Figure 8. ResNet-18 architecture summary

3.2.5. ResNet-50 deep learning model

ResNet-50 is considered to be of the latter category which has three layers deep residual blocks Figure 8. Figure 9 demonstrates the detailed layers and builds used to implement the ResNet-50 model.

Layers	Output Size	ResNet-50
conv1	112 x 112	7 x 7 , 64 , stride 2
conv2_x	56 x 56	3 x 3 max pool, stride 2
		$\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$
conv3_x	28 x 28	$\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 4$
		$\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 6$
conv4_x	14 x 14	$\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$
		$\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 6$
conv5_x	7 x 7	$\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$
		$\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 6$
	1 x 1	average pool, 100-d fully-connected, softmax
FLOPs		3.8x10 ⁹

Figure 9. ResNet-50 architecture summary

3.3. Classification

For decision-making, pattern recognition, and others, data classification is an important function for separating broad datasets into different classes [28]. A classification layer uses a fully connected layer and measures the cross-entropy loss in multi-class classification situations involving classes that are mutually exclusive. We customize the last layer of our model to accommodate a two-class classification.

Next, the model is assembled and the loss function is defined as binary cross-entropy, as the problem involves binary classification. The optimizer is configured to use the Adam optimizer, sometimes referred to as adaptive moment estimation. This optimizer is considered superior and more efficient than stochastic gradient descent (SGD) for training deep learning models. It is recognized as the most effective optimizer that has demonstrated improved precision [35].

4. RESULTS AND DISCUSSIONS

Approaching the performance improvement of colon cancer detection, in this section we discuss our results and how particular changes impact our results, from the process of reading the dataset to the classification of the result. Upon looking for the best CNN network to read and understand our dataset in this section we compared the results of the different CNN networks that we implemented prior and how they performed with our dataset, as shown in Table 3. The results presents that the CNN network ordered from the highest accuracy rate to the lowest is ResNet-50, followed by DenseNet-121, VGG-16, GoogLeNet, and ResNet-18 in that order. Moving forward, we will use ResNet-50 given the great results it has shown.

Table 3. Models performance trained on our dataset

Model	Accuracy
DenseNet-121	94.7%
VGG-16	92.4%
GoogLeNet	91.4%
ResNet-50	99.7%
ResNet-18	86.9%

After transforming the training dataset, we perform a train/test split, giving us two different sets: a training set and a testing set. We use this method to avoid underfitting but mostly overfitting the model. In order

to apply the model to new data, it first learns from the training set, which comprises known output. To ensure that our models are accurate, we have a test dataset or subsets available [36]. We implemented three different train/test ratios: 80 percent for training and 20 percent for testing, 70 percent for training and 30 percent for testing, and 60 percent for training and 40 percent for testing, and the results are as shown in Table 4. The results show minor variations, but we observe that the train/test split ratio that yields the maximum recall value is 80% training and 20% testing. Therefore, we will proceed with this ratio.

Table 4. Model performance with different training data to testing data ratios

Training data: testing data	Accuracy	F1-score	Recall	Precision	Specificity
80% : 20%	99.72%	99.7%	99.8%	99.7%	99.7%
70% : 30%	99.7%	99.7%	99.7%	99.7%	99.7%
60% : 40%	99.52%	99.5%	99.6%	99.5%	99.4%

In this paragraph, we provide the differences in results before and after applying SN, an image pre-processing approach widely employed with histopathology images on the dataset as illustrated in Table 5, from several performance perspectives. The use of the SN approach results in a notable enhancement in the overall performance of the model. Therefore, we will continue utilizing our stain-normalized dataset.

Table 5. Model performance with original images and with images after SN

Images	Accuracy	F1-score	Recall	Precision	Specificity
Original images	99.72%	99.7%	99.8%	99.7%	99.7%
Stain normalized images	99.88%	99%	99%	99.9%	99.9%

Furthermore, we analyze the outcomes obtained by employing 10, 20, and 30 epochs to train our model and evaluate the performance for each epoch count. An epoch is a parameter that determines the number of iterations the learning algorithm will traverse over the full training dataset. An era consists of one or more batches [37]. Upon conducting training sessions using 10, 20, and 30 epochs, the outcome is presented in Table 6. The results indicate that utilizing 30 epochs throughout the training phase yielded superior performance in comparison to 10 and 20 epochs.

Table 6. Model performance with different number of epochs

Number of epochs	Accuracy	F1-score	Recall	Precision	Specificity
10 epochs	99.83%	99.8%	99.7%	100%	100%
20 epochs	99.88%	99%	99%	99.9%	99.9%
30 epochs	99.94%	99.9%	100%	99.8%	99.8%

After comparing the results of the tests we conducted, we have chosen our final model as demonstrated in Figure 10. With ResNet-50 achieving the highest accuracy, we decided to use it as our feature extraction model moving forward. Later, after comparing cross-validation results, we found that the ratio of 80:20 training data to testing data achieved the best performance. Afterward, we compared the results before and after applying SN to the dataset images, and we found that the dataset with SN achieved better performance. For training, we compared the results of different numbers of epochs and found that using 30 epochs gave the best performance.

Deep learning models can be obscure in terms of their real-world performance. In this part, we thoroughly assess the performance of our ResNet-50 model. Figure 11 displays the graphical representation of the accuracy and loss. We utilized the Confusion Matrix to facilitate the assessment of our model. The Confusion Matrix is a structured representation of the outcomes of our predictions. It provides a summary of the accurate and inaccurate forecasts made for each category. The classification results on the test data using the ResNet-50 are depicted in Figure 12. The Confusion Matrix yielded the following values: true positive (TP) = 923, true negative (TN) = 876, false positive (FP) = 1, and false negative (FN) = 0. These values are utilized in the computation of F1-score, Recall, Precision, and Specificity. The F1-score achieved a result of

99.9%, recall achieved a perfect score of 100%, precision achieved a result of 99.8%, and specificity achieved a result of 99.8%.

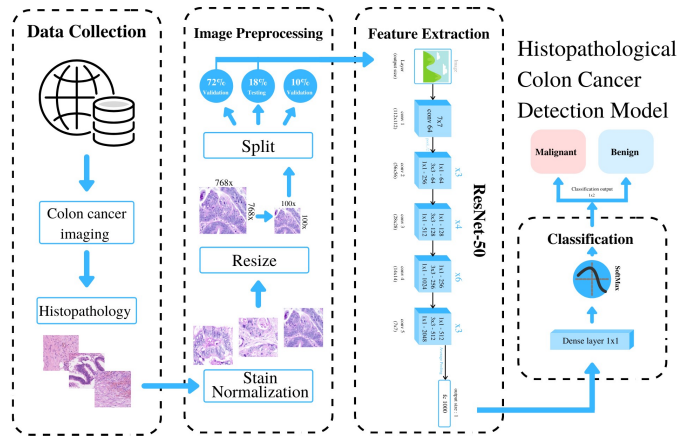


Figure 10. The final histopathological colon cancer detection model

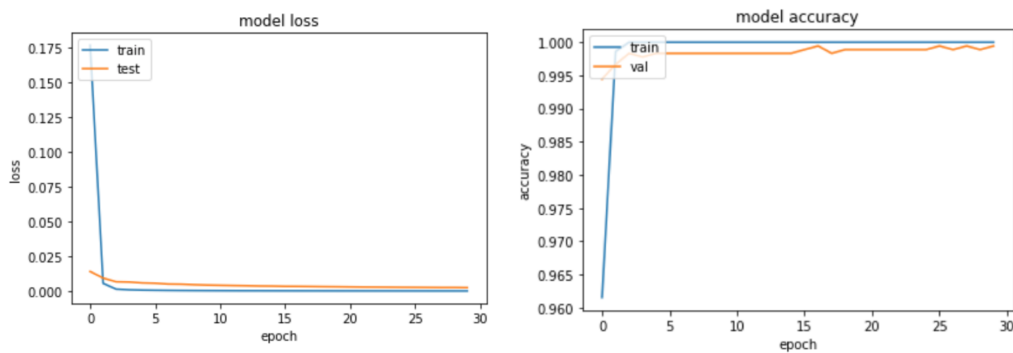


Figure 11. The ResNet-50 (with SN images) model accuracy and loss

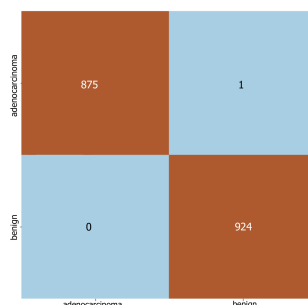


Figure 12. The confusion matrix of the ResNet-50 (with SN images) model results

Within this section, we conduct a comparative analysis between our suggested model and the prior research conducted. Previous studies that utilized deep learning models will be compared using the suggested ResNet-50 algorithm using SN images because it surpasses the other four models (as shown in Table 3). The measures for performance evaluation, including accuracy, F1-score, recall, accuracy, and specificity, are

displayed in Table 7. These metrics apply to both the existing deep learning models and the one we propose ResNet-50 model. The accuracy reached by our proposed ResNet-50 with SN images is 99.94% in comparison with DenseNet [6], GoogLeNet [7], Dual-CNN [8], ResNet-121 [9], Crafted-VGG [10], Resnet-50 [11], RCCNet [12], CNN+SampEn [14], and DenseNet-121 [15] which have 97.34%, 91.5%, 93.7%, 94%, 93.48%, 88%, 80.6%, 98.3%, and 98.53%, respectively, as represented on Figure 13.

Table 7. Analyzing the various models from prior research and comparing them to our suggested ResNet-50 (with SN pictures) model utilizing deep learning criteria for performance evaluation

Models	Accuracy (%)	F1-score (%)	Precision (%)	Recall (%)	Specificity (%)
DenseNet [6]	97.34	-	-	-	-
GoogLeNet [7]	91.5	82.9	-	-	-
Dual-CNN [8]	93.7	93.4	92.9	93.9	-
ResNet-121 [9]	94	-	-	-	-
Crafted-VGG [10]	93.48	-	-	95.1	92.76
Resnet-50 [11]	88	-	-	93	83
RCCNet [12]	80.6	78.87	-	-	-
GLCM [13]	-	89.55	96.67	83.33	-
CNN+SampEn [14]	98.3	-	-	-	-
DenseNet-121 [15]	98.53	-	-	98.63	98.48
ResNet-50 with SN images (our proposed)	99.94	99.99	99.8	100	99.8

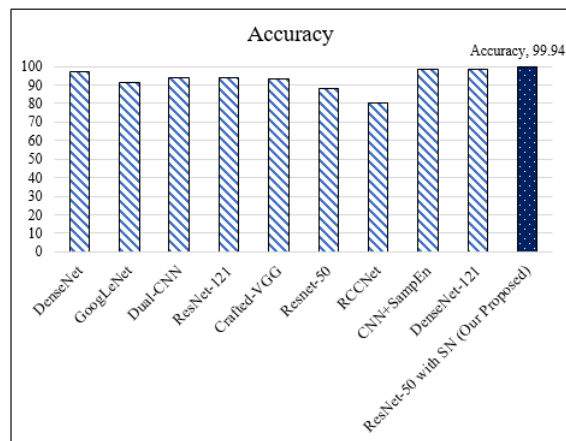


Figure 13. Analyzing the various models from prior research and comparing their accuracy performance metrics (DenseNet [6], GoogLeNet [7], Dual-CNN [8], ResNet-121 [9], Crafted-VGG [10], Resnet-50 [11], RCCNet [12], CNN+SampEn [14], and DenseNet-121 [15]) and our proposed ResNet-50 with SN images model

Figure 14 shows the F1-score reached by our proposed ResNet-50 with SN images compared with the previous studies. In Figure 14, the F1-score achieved by our proposed model is 99.9% in comparison with GoogLeNet [7], Dual-CNN [8], RCCNet [12], and GLCM [13], which have 82.9%, 93.4%, 78.87%, and 89.55%, respectively. while our proposed model reached 99.8% of precision in comparison with Dual-CNN [8] with 92.9% and GLCM [13] with 96.67%.

The recall performance metric of our proposed ResNet-50 model obtained 100% in comparison with Dual-CNN [8], Crafted-VGG [10], Resnet-50 [11], GLCM [13], and DenseNet-121 [15], which have 93.9%, 95.1%, 93%, 83.33%, and 98.63%, respectively. While the specificity achieved by our ResNet-50 model is 99.8% compared with crafted-VGG [10], Resnet-50 [11], and DenseNet-121 [15], which have 92.76%, 83%, and 98.48%, respectively. The recall is presented in Figure 15, while the specificity is shown in Figure 16. Our proposed ResNet-50 was able to do this because the features were better represented in the stain-normalized images. This let the deep learning model find the correlation between the data's input features, which led to better prediction abilities.

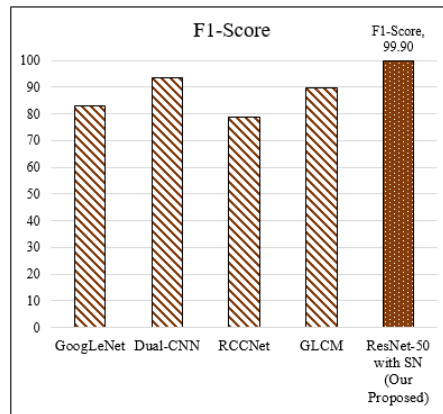


Figure 14. Analyzing the various models from prior research and comparing their F1-score performance metrics (GoogLeNet [7], Dual-CNN [8], RCCNet [12], and GLCM [13]) and our proposed ResNet-50 with SN images model

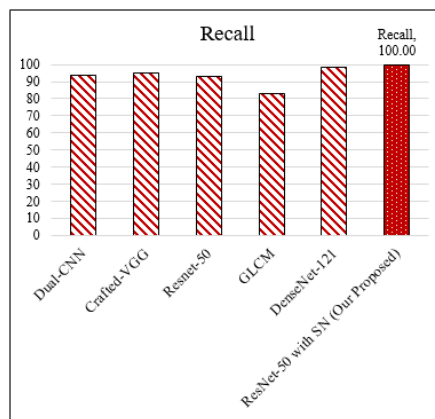


Figure 15. Analyzing the various models from prior research and comparing their recall performance metrics (Dual-CNN [8], Crafted-VGG [10], Resnet-50 [11], GLCM [13], and DenseNet-121 [15]) and our proposed ResNet-50 with SN images model

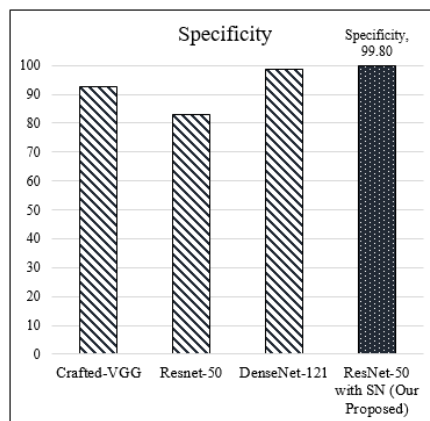


Figure 16. Analyzing the various models from prior research and comparing their specificity performance metrics (Crafted-VGG [10], Resnet-50 [11], and DenseNet-121 [15]) and our proposed ResNet-50 with SN images model

5. CONCLUSION

Throughout this study, we have explored many topics in the field of cancer detection, more specifically histopathological cancer detection. After we defined our topic of interest as the colon, we researched ways to improve the performance of the detection model by examining the effect of SN of histopathology images on the performance of CNN models. To enhance the histological colon cancer detection model, we initiate the process by utilizing five distinct deep-learning models: DenseNet-121, VGG-16, GoogLeNet, ResNet-50, and ResNet-18. We also investigate the impact of SN on histopathology images and its enhancement of our model's performance. According to the results, ResNet-50 had the greatest accuracy rate of 99.7% compared to the other four models. Furthermore, when utilizing the ResNet-50 model, we discovered that the optimal ratio for training and testing data is 80% and 20%, respectively. Additionally, we conducted an experiment using the ResNet-50 model on both the original images and the stain-normalized images from the same dataset. Our findings indicate that the stain-normalized images yielded a higher accuracy rate of 99.94% when trained for 30 epochs. We conducted a comparative analysis of nine different models that were previously studied. We evaluated their accuracy, F1-score, precision, recall, and specificity. The results indicate that our suggested ResNet-50 model, which uses stain normalized images, surpasses all the other models that were examined in the prior studies. Consequently, we enhanced the overall efficacy of the histopathological colon cancer detection model, exhibiting notable outcomes in contrast to earlier investigations in that domain.

ACKNOWLEDGEMENTS

The researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2024-9/1)




REFERENCES

- [1] World Health Organisation: Cancer, "World Health Organization," *Public Health Economics*, vol. 6, no. 7, pp. 527–528, Jul. 1949, doi: 10.1177/107755874900600710.
- [2] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, Nov. 2018, doi: 10.3322/caac.21492.
- [3] R. Khatun and S. Chatterjee, "Machine learning approach for segmenting glands in colon histology images using local intensity and texture features," in *Proceedings of the 8th International Advance Computing Conference, IACC 2018*, Dec. 2018, pp. 314–320, doi: 10.1109/IADCC.2018.8692135.
- [4] K. A. H. Sheriff, "Colon cancer – a review," *Journal of Pharmaceutical Sciences and Research*, vol. 8, no. 10, pp. 1191–1194, 2016.
- [5] I. Pacal, D. Karaboga, A. Basturk, B. Akay, and U. Nalbantoglu, "A comprehensive review of deep learning in colon cancer," *Computers in Biology and Medicine*, vol. 126, p. 104003, Nov. 2020, doi: 10.1016/j.combiomed.2020.104003.
- [6] J. Ke, Y. Shen, Y. Guo, J. D. Wright, N. Jing, and X. Liang, "A high-throughput tumor location system with deep learning for colorectal cancer histopathology image," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12299 LNAI, 2020, pp. 260–269.
- [7] R. Awan, S. Al-Maadeed, R. Al-Saady, and A. Bouridane, "Glandular structure-guided classification of microscopic colorectal images using deep learning," *Computers and Electrical Engineering*, vol. 85, p. 106450, Jul. 2020, doi: 10.1016/j.compeleceng.2019.106450.
- [8] S. C. Kosaraju, J. Hao, H. M. Koh, and M. Kang, "Deep-Hipo: multi-scale receptive field deep learning for histopathological image analysis," *Methods*, vol. 179, pp. 3–13, Jul. 2020, doi: 10.1016/j.ymeth.2020.05.012.
- [9] N. Dif and Z. Elberichi, "A new deep learning model selection method for colorectal cancer classification," *International Journal of Swarm Intelligence Research*, vol. 11, no. 3, pp. 72–88, Jul. 2020, doi: 10.4018/IJSIR.2020070105.
- [10] H. Yoon *et al.*, "Tumor identification in colorectal histology images using a convolutional neural network," *Journal of Digital Imaging*, vol. 32, no. 1, pp. 131–140, Feb. 2019, doi: 10.1007/s10278-018-0112-9.
- [11] D. Sarwinda, R. H. Paradisa, A. Bustamam, and P. Anggia, "Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer," *Procedia Computer Science*, vol. 179, pp. 423–431, 2021, doi: 10.1016/j.procs.2021.01.025.
- [12] S. H. S. Basha, S. Ghosh, K. K. Babu, S. R. Dubey, V. Pulabaigari, and S. Mukherjee, "RCCNet: an efficient convolutional neural network for histological routine colon cancer nuclei classification," in *2018 15th International Conference on Control, Automation, Robotics and Vision, ICARCV 2018*, Nov. 2018, pp. 1222–1227, doi: 10.1109/ICARCV.2018.8581147.
- [13] L. Jiao, Q. Chen, S. Li, and Y. Xu, "Colon cancer detection using whole slide histopathological images," in *IFMBE Proceedings*, vol. 39 IFMBE, 2013, pp. 1283–1286.
- [14] L. F. Segato dos Santos, L. A. Neves, G. B. Rozendo, M. G. Ribeiro, M. Zanchetta do Nascimento, and T. A. A. Tosta, "Multidimensional and fuzzy sample entropy (SampEnMF) for quantifying H&E histological images of colorectal cancer," *Computers in Biology and Medicine*, vol. 103, pp. 148–160, 2018, doi: 10.1016/j.combiomed.2018.10.013.
- [15] D. Sarwinda, A. Bustamam, R. H. Paradisa, T. Argyadiva, and W. Mangunwardoyo, "Analysis of deep feature extraction for colorectal cancer detection," in *2020 4th International Conference on Informatics and Computational Sciences (ICICoS)*, Nov. 2020, pp. 1–5, doi: 10.1109/ICICoS51170.2020.9298990.




- [16] A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wilson, L. A. DeLand, and S. M. Mastorides, "Lung and colon cancer histopathological image dataset (lc25000)," arXiv preprint, 2019.
- [17] K. Tomczak, P. Czerwińska, and M. Wiznerowicz, "The cancer genome atlas (TCGA): an immeasurable source of knowledge," *Współczesna Onkologia*, vol. 1A, pp. A68–A77, 2015, doi: 10.5114/wo.2014.47136.
- [18] J. N. Kather et al., "Multi-class texture analysis in colorectal cancer histology," *Scientific Reports*, vol. 6, 2016, doi: 10.1038/srep27988.
- [19] K. Sirinukunwattana, D. R. J. Snead, and N. M. Rajpoot, "A stochastic polygons model for glandular structures in colon histology images," *IEEE Transactions on Medical Imaging*, vol. 34, no. 11, pp. 2366–2378, 2015, doi: 10.1109/TMI.2015.2433900.
- [20] K. Sirinukunwattana, S. E. A. Raza, Y. W. Tsang, D. R. J. 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, doi: 10.1109/TMI.2016.2525803.
- [21] K. Sirinukunwattana et al., "Gland segmentation in colon histology images: the glas challenge contest," *Medical Image Analysis*, vol. 35, pp. 489–502, Jan. 2017, doi: 10.1016/j.media.2016.08.008.
- [22] F. Ciompi et al., "The importance of stain normalization in colorectal tissue classification with convolutional networks," in *Proceedings - International Symposium on Biomedical Imaging*, Apr. 2017, pp. 160–163, doi: 10.1109/ISBI.2017.7950492.
- [23] Z. Gandomkar, P. C. Brennan, and C. Mello-Thoms, "MuDeRN: multi-category classification of breast histopathological image using deep residual networks," *Artificial Intelligence in Medicine*, vol. 88, pp. 14–24, Jun. 2018, doi: 10.1016/j.artmed.2018.04.005.
- [24] J. N. Kather et al., "Multi-class texture analysis in colorectal cancer histology," *Scientific Reports*, vol. 6, no. 1, p. 27988, Jun. 2016, doi: 10.1038/srep27988.
- [25] B. E. Bejnordi et al., "Stain specific standardization of whole-slide histopathological images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 2, pp. 404–415, Feb. 2016, doi: 10.1109/TMI.2015.2476509.
- [26] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.
- [27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
- [28] Q. A. Al-Hajja and A. Adebajo, "Breast cancer diagnosis in histopathological images using ResNet-50 convolutional neural network," in *IEMTRONICS 2020 - International IOT, Electronics and Mechatronics Conference, Proceedings*, Sep. 2020, pp. 1–7, doi: 10.1109/IEMTRONICS51293.2020.9216455.
- [29] S. Nemade and S. Sonavane, "Comparative analysis of geometric transformation effects for image annotation using various CNN models," in *Advances in Intelligent Systems and Computing*, vol. 1155, 2020, pp. 362–369.
- [30] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.
- [31] A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, Dec. 2020, doi: 10.1007/s10462-020-09825-6.
- [32] S. Karen and Z. Andrew, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, p. 14, 2015.
- [33] D. R. D. Darmawan, D. Wahyudin, D. Rahadian, A. Suryadi, "Deep learning methods for anatomical landmark detection in video capsule endoscopy images," *Proceedings of the Future Technologies Conference*, vol. 3, 2022.
- [34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, vol. 2016-Decem, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [35] S. Doshi, "Various optimization algorithms for training neural network," *Towards data science*, 2019. <https://towardsdatascience.com/optimizers-for-training-neural-network-59450d71caf6>.
- [36] A. Bronshtein, "Train/test split and cross validation in Python," *Towardsdatascience.Com*, 2017. www.towardsdatascience.com/train-test-split-and-cross-validation-in-python-80b61beca4b6.
- [37] J. Brownlee, "What is the difference between a batch and an epoch in a neural network?," *Machine Learning Mastery*, no. July, pp. 3–4, 2018.

BIOGRAPHIES OF AUTHORS






Dina M. Ibrahim    has been an associate professor at the Information Technology Department, College of Computer, Qassim University, KSA, since 2015. In addition, Dina works as an assistant professor at the Computers and Control Engineering Department, Faculty of Engineering, Tanta University, Egypt. She was born in the United Arab Emirates, and her B.Sc., M.Sc., and Ph.D. degrees were obtained from the Computers and Control Engineering Department, Faculty of Engineering, Tanta University, in 2002, 2008, and 2014, respectively. Dina worked as a consultant engineer, database administrator, and vice manager on the Management Information Systems (MIS) Project, at Tanta University, Egypt, from 2008 until 2014. Her research interests include networking, wireless communications, machine learning, security, and the Internet of Things. Dina has published about 70 articles in various refereed international journals and conferences. She has been serving as a reviewer in the Wireless Network (WINE) Journal since 2015. Dina has also served as a co-chair of the International Technical Committee for the Middle East Region of the ICCMIT conference since 2020. She can be contacted at email: d.hussein@qu.edu.sa.



Mohammad Ali A. Hammoudeh    is an Assistant Professor at the Information Technology Department, Collage of Computer (CoC), Qassim University, Saudi Arabia (KSA) from 2010 till now. In addition, Mohammad has worked as an assistant professor at the Electrical Engineering Department, College of Engineering, ISRA University-Jordan. He was born in Kuwait, and his B.Sc., M.Sc., and Ph.D. degrees were taken from the National Technical University NTUU “KPI”– Kyiv (Ukraine), in 2000, 2002, and 2007, respectively. Mohammad works as a member of the information technology group of e-learning research (TIGER), ISRA University-Jordan, from 2008 until 2010, then as a deputy in the CoC scientific research committee, Qassim University (QU), Saudi Arabia. His research interests include Computing and Networks, e-learning systems, IoT, Mobile 5G, Security, Renewable and Sustainable Energy, and Cloud Technology. Mohammad has published more than 20 articles in various refereed international journals and conferences. Also, he has served as a reviewer in IEEE, MDPI, ACM, Wiley, Hindawi, and in Information scientific research deanship (SRD), QU since 2015. He can be contacted at email: maah37@qu.edu.sa.



Tahani M. Allam    works as a lecturer at Computers and Control Engineering Department, Faculty of Engineering, Tanta University, Egypt. She was born in Kuwait, and her B.Sc., M.Sc., and Ph.D. degrees were taken from the Computers and Control Engineering Department, Faculty of Engineering, Tanta University in 2002, 20010, and 2018, respectively. Tahani has worked as a Consultant Engineer, and then a General Supervisor of the Human Resources Management Program on the Management Information Systems (MIS) Project- at Tanta University since 2010 till now. She has been the Director of the Job Performance Evaluation Unit at Tanta University since 2021. Her research interests include cloud computing, Semantic Web, machine learning, security, and the Internet of Things. Tahani has published about 10 articles in various refereed international journals and conferences. She can be contacted at email: tahany@f-eng.tanta.edu.eg.