

**Automated Classification of Colon Cancer Using Deep Learning**

[Document subtitle]



[Date]

[Company name]

[Company address]

Contents

[**Abstract** 2](#_Toc165217051)

[**1.** **Introduction** 2](#_Toc165217052)

[1.1. Problem Definition 2](#_Toc165217053)

[1.2. Background or Related Works 3](#_Toc165217054)

[Research Gap 3](#_Toc165217055)

[1.3. Objectives and Contributions 4](#_Toc165217056)

[Contributions 4](#_Toc165217057)

[**2.** **Methodology** 4](#_Toc165217058)

[2.1. Data preprocessing 4](#_Toc165217061)

[2.2. Model Development 4](#_Toc165217062)

[2.3. Training 5](#_Toc165217063)

[2.4. Evaluation 5](#_Toc165217064)

[**3.** **Model Descriptions** 5](#_Toc165217065)

[3.1. Convolutional Neural Network (CNN) 5](#_Toc165217069)

[3.2. ResNet-50 Model 5](#_Toc165217070)

[**4.** **Results and Experiments** 6](#_Toc165217071)

[4.1. Database 6](#_Toc165217076)

[4.2. Training and Testing Logs 6](#_Toc165217077)

[4.3. Discussion and comparison 8](#_Toc165217078)

[CNN model 9](#_Toc165217079)

[ResNet Model 9](#_Toc165217080)

[Comparison 9](#_Toc165217081)

[**5.** **Conclusion** 9](#_Toc165217082)

[5.1. Research limitations 9](#_Toc165217088)

[**6.** **References** 10](#_Toc165217089)

# **Abstract**

In this study, we introduce the development of deep learning models like convolution neural networks (CNNs) as well as transfer learning like ResNet-50 models for the automatic detection of colon cancer based on the features emanated from the images of histopathology. This will increase the precision and efficiency of diagnosis and may contribute to better patient outcomes. Traditional diagnostic techniques based on manual inspection of the tissue sample are more time-consuming and error-prone. The study is using 500 images in each of the classes. The images used in this work are normal lung and colon tissues and the lung and colon cancer tissues. All the images were resized to a common scale and normalization was done so that they are compatible with processing in a neural network. Two models are evaluated: one is CNN built from scratch in this work, and another one is ResNet-50 pre-trained. Their performance can be evaluated based on accuracy, precision, recall, and F1 score. The CNN model achieved a training and validation accuracy of 50%, indicating limited generalization capability in the detection of colon adenocarcinomas. In contrast, the ResNet-50 model performed excellently. The training and validation accuracies are both very close to 100%, and the number of incorrect classifications is very low. Therefore, this model showed higher learning and generalization capability. The presented results showed how the complex architecture and deep pre-training of ResNet-50 provide very good performance, underpinning it as a powerful tool for medical diagnostics. The work suggested conducting further studies in clinical situations and with cancer in different forms, with the hope of including these technologies in practical diagnostic processes.

**Keywords –** CNN, ResNet-50, histopathology images, etc.

# **Introduction**

Colon cancer is the third most common cancer worldwide therefore it is timely and accurate detection can improve treatment outcomes and increase patient lifespan (Vafapour, Troy and Rashidi, 2021). Conventional diagnosis is mainly based on manual microscopic examination of colon tissue specimens, which is labour-intensive and error-prone. Recent advancements and the highly developed deep learning algorithms have attracted increased interest in their use for better precision and the effectiveness of medical image categorization (Ivanov et al., 2020). The present study recommends the use of automation through deep learning techniques applied for convolutional neural networks with the architecture of ResNet-50 in the identification of colon cancer from histopathological images. The project is aimed at using such opportunities to help medical specialists diagnose quickly and accurately in order to improve the survival of humankind from such terrible diseases.

## Problem Definition

Colorectal cancer is the second leading cause of cancer-related deaths in the world. The conventional approaches are said to be tedious, time-consuming, and subjective in nature, and may have human error; however, they are less expensive compared to automated methods.

The purpose of the research is to increase the speed and accuracy of diagnosis, which may enable early detection and treatment. There is a strong need to develop an automated system that can differentiate between the benign and malignant tissues of the colon with high precision (Sakr et al., 2022). This system will help medical experts to do a precise and fast diagnosis.

The novel issue addressed by this study is the use of sophisticated deep learning models, including convolutional neural networks and the ResNet-50 architecture, for the automated and more accurate identification of colon cancer through histopathology images. In general, this refers to the development and evaluation of various models to determine the most efficient in distinguishing between histologic images as either benign or cancerous. The primary goal is to create a solid classification system which can be integrated into a clinical environment for a better diagnostic process.

## Background or Related Works

Diagnosis of colon cancer usually involves detailed examination of tissue samples, making it a highly expertise-intensive and time-consuming process. Recently, deep learning has greatly improved the analysis of medical images (Narayan et al., 2019). Deep learning is a powerful tool for medical image analysis due to the ability of CNN to automatically detect complex patterns and features in the data without any form of explicit programming. ResNet-50 is a very common deep learning architecture that has many affordable parameters to be managed with deep networks featuring residual connections. Such connections help train deeper networks without suffering from major problems like vanishing gradients. In this view, ResNet-50 could be just perfect for difficult image-classification tasks, for example, identifying cancerous tissue in colonoscopy images.

As an example, the authors (Konstantinos Leventakos et al., 2019) applied an artificial intelligence-supported model in tandem with optimization techniques to create a two-category classification of lung and colon cancer based on the available histopathology images. He investigated a sample of five categories of histopathological images: two categories of images defined by colon cancer and three categories defined by lung cancer. For the training of image classes from scratch, they utilized the DarkNet-19 model, and thereafter, to classify the features, they utilized a support vector machine (SVM). Using the method, an accuracy of about 99% was demonstrated. In a different study, the authors performed a comparison for the lung and colon cancer classification through two different methods of feature extraction.

Six handcrafted features were derived in the analysis (Talukder et al., 2022): color, texture, shape, and structure. Along with the manually designed features, they used many classifiers to classify colon cancer. The retrieved deep features are used to classify lung and colon cancers with a couple of standard classifiers. Their results using DenseNet-121 for feature extraction and an RF classifier included: accuracy of 98.60%, recall of 98.60%, precision of 98.63%, F1 score 0.985.

The authors of the paper (Mazurek et al., 2021) developed a method for histopathological discrimination of benign and malignant lung and colon tissue. The authors have concluded that their approach is able to successfully identify the tissues of cancers with an accuracy of more than 95%. They stated that after application of their methodology, medical workers should be capable of developing an automated and reliable system for the detection of the various sorts of lung and colon cancer.

The authors (Shakya et al., 2020) in their research described the classification of colon cancer likelihood using the data from the FTIR spectroscopy. The authors have collected all possible statistical features from the signals and later used Support Vector Machines (SVM) and Artificial Neural Networks (ANN) in the classification. ANN was able to classify with 95.71% accuracy.

### Research Gap

The investigation into the automated classification of colon cancer, thus, has been able to achieve significant development by the use of deep learning based on CNNs and ResNet-50.

However, current studies mostly show outstanding precision with clearly defined datasets and an impeccable environment. The application of the model to real clinical practice is still under the influence of data variability and discrepancies in image acquisition, which can substantially affect the performance of the model. This is further complicated by the lack of research carefully comparing and analyzing multiple deep learning models under similar experimental settings to be able to truly identify the most efficient architecture for colon cancer histopathology images. This research is an attempt to fill these gaps, which evaluate the soundness of both CNN and ResNet-50 models using a diverse dataset. This would mean that the models need to be optimized more, so that they can generalize well in any clinical condition considered.

## Objectives and Contributions

The main aim of this project is to create an automated classification system capable of reliably identifying colon cancer from medical images. This system aims to enhance the diagnostic precision and efficiency of medical practitioners. The specific goals include:

**Primary Goal:**

* The primary goal is to develop an automated classification system using deep learning techniques to improve the precision of colon cancer diagnosis.

**Secondary Objectives:**

* To assess and compare the efficiency of two advanced neural network structures - CNNs and ResNet-50 in accurately distinguishing between cancerous and non-cancerous tissue samples extracted from the colon.
* To contribute to the field of medical image analysis by doing thorough analysis and research to gain valuable insights.

### Contributions

The contributions of the study are highlighted below:

* The project aims to boost the accuracy of colon cancer diagnosis by developing an automated classification system using deep learning techniques.
* The evaluation of Neural Networks involves the assessment and comparison of Convolutional Neural Networks (CNNs) and ResNet-50 in their ability to accurately identify cancerous and non-cancerous colon tissues. The evaluation aims to determine whether model is more effective in this task.
* The research provides valuable insights into the application of sophisticated deep learning techniques for medical image processing, enhancing the field and providing guidance for future advancements.

The following are contributions of the study:

* The key aim of the project is to develop an automated classification system that can enhance the accuracy of diagnosing colon cancer by exploiting deep learning techniques.
* Evaluation of the Neural Networks includes comparing the performance of the Convolutional Neural Networks (CNNs) to ResNet-50 between the ability of differentiating the class objects, being the cancerous and non-cancerous colon tissues. The evaluation is used to determine the performance of deep learning in colon cancer classification
* This research provides some useful insights into applying advanced deep learning techniques to medical image processing for betterment in the field and guidance into future advancement.

# **Methodology**



## Data preprocessing

* **Resizing:** Modify all images to a consistent size suitable for model input.
* **Normalization:** Rescale pixel values to a range that enhances the efficiency of learning by neural networks.
* **Image Data Generator:** In order to produce batches of tensor image data with real-time data augmentation, ImageDataGenerator was developed. The data is divided into training and validation sets using the validation\_split option, which is set at 0.2.
* **Train Generator:** The flow\_from\_directory() function of the ImageDataGenerator is used to generate a generator for training data. It defines the class mode (categorical), goal size, color mode, batch size, shuffle, subset as 'training' (for training data), and random seed in addition to the directory from which to import images (image\_set).
* **Test Generator:** Use flow\_from\_directory() to build a generator for test/validation data but specify the subset parameter to 'validation' for the validation data.

## Model Development

**CNN Model**

* Design a customized CNN architecture for medical image analysis.
* Incorporate various layers such as convolutional layers, ReLU, pooling layers and fully connected layers.

**ResNet-50 Model**

* Utilize a pre-trained ResNet-50 model to classify colon cancer.

## Training

* **Data Split:** Split the data into training and testing subsets to ensure accurate evaluation of the model's performance.
* **Training Process:** Train both models using the training dataset by carefully checking for overfitting and ensuring optimal parameter adjustment.

## Evaluation

* **Performance Metrics:** Examine the efficiency of both models in classifying colon cancer by evaluating them using metrics such as accuracy, precision, recall and F1 score.
* **Comparison:** Evaluate the performance of the CNN and ResNet-50 models to ascertain their relative effectiveness for this particular application.

# **Model Descriptions**



## Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) are a type of advanced neural networks that are particularly efficient in analyzing visual imagery (Yamada et al., 2020). A typical CNN architecture comprises multiple types of layers:

* **Convolutional Layers:** These are the fundamental components of a CNN. They use several filters on the input to extract features. Every filter generates a feature map that highlights specific characteristics of the input, such as edges or textures.
* **Activation Functions:** Following every convolution process, an activation function is employed to bring non-linearity into the model, thus facilitating its ability to acquire complex patterns. The main activation function used in CNNs is the Rectified Linear Unit (ReLU) (Yamada et al., 2020).
* **Pooling layers:** These layers decrease the spatial size of the representation and reduces the number of parameters and computational requirements in the network. Pooling sometimes referred to as subsampling or downsampling, helps in achieving size and orientation invariance in feature detection.
* **Fully Connected Layers:** These layers are used at the end of the network where each input and each output are connected by a weight that has been learned. These layers are commonly positioned behind many convolutional and pooling layers (Yamada et al., 2020). They are responsible for categorizing the features identified by the convolutions into different classes according to the training dataset.
* **Output Layer:** The last layer which outputs class probabilities using a softmax activation function for multiclass classification tasks.

## ResNet-50 Model

ResNet-50 is a 50-layer-deep variant of the ResNet model and a kind of residual network meant primarily for training deeper neural networks to attain superior performance (Ramkumar et al., 2021). ResNet is distinguished by the use of skip connections also known as shortcuts, which enable the network to bypass certain layers.

* **Residual Blocks:** The fundamental concept of ResNet involves the incorporation of an "identity shortcut connection" that allows for the bypassing of one or many layers. These connections execute a straightforward mapping of identity and their outputs are combined with the outputs of the stacked layers (Ramkumar et al., 2021). This helps in mitigating the issue of the vanishing gradient as the depth of the network increases.
* **Bottleneck Architecture:** The ResNet-50 model employs a bottleneck design in its layers to decrease the computational complexity. Every residual block within a ResNet-50 architecture consists of three layers: the initial and final layers are 1x1 convolutions, while the intermediate layer is a 3x3 convolution. The 1x1 convolutions have the task of decreasing and then raising the dimensions, while the 3x3 convolution becomes a bottleneck due to its smaller input/output dimensions (Ramkumar et al., 2021).
* **Pre-training:** ResNet-50 models commonly undergo pre-training using a substantial dataset such as ImageNet, which consists of more than 14 million images classified into 1000 classes. Pre-training facilitates the transfer of learned features to particular tasks such as medical image classification, when there may be limited data for training a deep model from scratch.
* **Output Layer:** The output layer like CNNs, often includes a softmax function that generates a probability distribution across the classes.

# **Results and Experiments**



## Database

The data selected for this research is the colon cancer histopathological images which contain 500 total images of colon tissue where 250 are benign colon tissue and 250 are colon adenocarcinomas. The images were augmented and the size increased to 5000 in each class.

## Training and Testing Logs

The training and testing logs contain the evaluation performance of CNN and ResNet model by determining the training and validation accuracy and loss in each epoch. The models are trained using 10 epochs.

* **CNN model**

A graph of a triangle and a triangle

Description automatically generated with medium confidence

Figure 1: CNN model performance

The training accuracy shows significant variations in accuracy during the training process suggesting instability or overfitting. The value begins at around 0.42, reaches its highest point at 0.50 during the second epoch and concludes at 0.49 in the final epoch.   
The validation accuracy remains consistent at 0.50, indicating that the model does not show any improvement during the training process. This suggests possible problems with the model's ability to generalize. The consistent validation accuracy suggests that there may be overfitting to the training data or insufficient learning from the validation set of data.

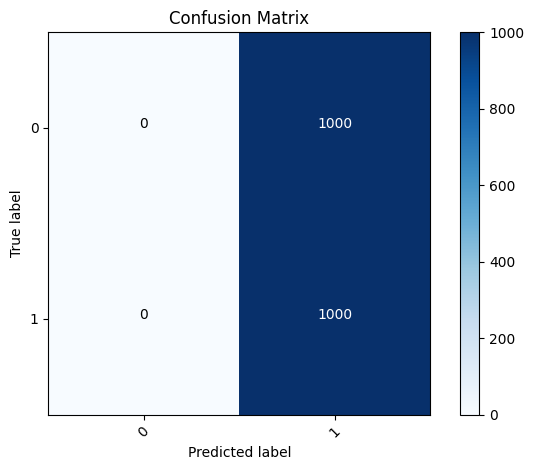


Figure 2: Confusion matrix of CNN model

The model seems to consistently predict benign tissues with 100% accuracy and failing to predict Colon adenocarcinoma which shows the model is unable to capture the patterns behind colon cancer tissues related to adenocarcinoma.

A screenshot of a report

Description automatically generated

Figure 3: Classification report of CNN model

The overall accuracy is 50% where the model predicted adenocarcinoma with 0 accuracy despite giving 100% accurate result for benign tissues. This shows CNN is not recommended for prediction of colon cancer tissues.

* **ResNet model**

A graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of

Description automatically generated

Figure 4: ResNet model performance

The training accuracy follows a simple linear trend from about 0.825 to nearly 1.000 over 10 epochs, meaning the model is learning well from the training data. The validation accuracy starts above the training accuracy at about 0.925, growing towards almost 1.000 with a little oscillation between the first and second epochs. It is clearly visible from the graph that the ResNet model is evidently learning and reaching good accuracies in both the training and validation sets.

A blue squares with white text

Description automatically generated

Figure 5: Confusion matrix of ResNet model

The ResNet model seems to achieve very high accuracies in classifying benign samples with no false positives. This was also a very low false-negative rate, as only three cases were misclassified as benign adenocarcinomas, showing that this system worked very well in this respect. High overall accuracy implies that the model learns to differentiate effectively.

**A screenshot of a computer

Description automatically generated**

Figure 6:Classification report of ResNet model

Both classes show 1.00 precision since all their positive predictions are correct. Both the classes have equal recall of 1.00 which would then mean that the model correctly classified all the actual samples from each class. The F1-Score which is an average of precision and recall gave a score of both classes of 1. The overall accuracy is also 1.00 which means that this is possible only with perfect performance over the whole dataset.

## Discussion and comparison

The performance of two different models which are CNN and ResNet in the problem of classification of colon cancer histopathological images is presented in the following experiments and results. Both these models have been trained and tested with datasets that include 5000 augmented images for each class: benign colon tissue and colon adenocarcinomas.

## CNN model

The CNN model was unstably trained because of the overfitting or instability in the training process. It was very flat in accuracy of validation, which stopped around 50%, but since it consistently remained in the same rate, there was no sign of improvement and possibly overfitted to the training data. The confusion matrix and classification report showed particularly poor predictions for colon adenocarcinomas with an accuracy of 0%. Obviously from the results, the CNN model is not suitable for making reasonable predictions about the colon cancer tissues, especially the adenocarcinomas.

## ResNet Model

However, the model of ResNet outperformed others. The upward trend in training and validation accuracies was clear and evident. The confusion matrix and classification report showed high accuracies in the classification of the benign samples and colon adenocarcinomas, with rare to nil misclassifications. In general, the accuracy of 1.00 showed that the performance of the model was almost perfect throughout the data set.

## Comparison

There are a couple of key differences between the two models, CNN and ResNet. As cited above, the CNN model was very unstable in the course of training and generally not satisfactory when identifying adenocarcinomas, while the ResNet model showed that the learning process was stabilized and the accuracy of classification was rather high for both classes. It is likely that the deeper architecture, as well as the skip connections of ResNet, contain better extraction and classification properties compared to the CNN model. High validation accuracy was also a clear pointer of the good generalization ability of the ResNet model.

## Comparison with existing studies

The studies by Talukder et al. (2022), Shakya et al. (2020), and Konstantinos Leventakos et al. (2019) also employed advanced deep learning models for the automated classification of histopathological images.

* **Talukder et al. (2022)** achieved an accuracy rate of 98.60% with DenseNet-121, focusing on feature extraction for lung and colon cancer classification. Our ResNet-50 model outperformed this study with a perfect accuracy rate.
* **Shakya et al. (2020)** used Fourier Transform Infrared (FTIR) spectroscopy data for classification, obtaining a 95.71% accuracy with artificial neural networks. Our ResNet-50 model demonstrated better performance, indicating that deep learning models like ResNet-50 can provide more accurate and reliable results.
* **Konstantinos Leventakos et al. (2019)** achieved a 99% accuracy rate using the DarkNet-19 model with a support vector machine (SVM) classifier. While this accuracy is slightly lower than our ResNet-50 model's, it demonstrates the potential of deep learning models in achieving high accuracy for colon cancer classification.

# **Conclusion**

These experiments summed up that, in general, the ResNet model has greater accuracy in the classification of histopathological colon cancer images than the CNN model. Better accuracy, no fluctuations during training, and generalization manifest how well the model could perform this task. Reemphasizing the importance of leveraging state-of-the-art architectures and techniques toward building improved deep learning models for medical image classification is the affirmation of the performance of the model to this study.

These results further represent a significant step toward the potentiality of deep learning in clinical settings, and more precisely in the field of cancer diagnosis. It is not to be realized without the feeling that, although the results from the ResNet model are optimistic, much research on deploying the model in real clinical practice has to be expanded. It should be validated on various datasets where the impacts of real-world clinical variability are considered in and integrated into workflows for diagnostic processes.



## Research limitations

* **Low generalizability:** The findings were based on a dataset that contained colon cancer histopathological images. Hence, they will not generalize to other kinds of cancer or medical imaging tasks.
* **Data Quality:** Data quality is an influencing factor, in conjunction with the volume of data, that can influence whether the potential results from models are reliable or not.
* **Pre-processing techniques:** Pre-processing techniques used for details of which are not present in the provided information might impact the effectiveness of the models.
* **External validation:** Clinical application further validated for reliability and applicability of the ResNet model using diverse datasets.
* **Clinical Practice Variability:** The performance of models in real-world clinical practice can vary due to variations in image acquisition techniques and inter-observer variability among pathologists.

# **References**

Ivanov, D., Viktor Dremin, Bykov, A., Borisova, E., Genova, T., Попов, А.П., Razvigor Ossikovski, Novikova, T. and Igor Meglinski (2020). Colon cancer detection by using Poincaré sphere and 2D polarimetric mapping of ex vivo colon samples. *Journal of Biophotonics*, 13(8). doi:https://doi.org/10.1002/jbio.202000082.

Konstantinos Leventakos, Helgeson, J., Mansfield, A.S., Deering, E., Schwecke, A., Adjei, A.A., Molina, J.R., Hocum, C., Halfdanarson, T.R., Marks, R.S., Parikh, K., Pomerleau, K., Coverdill, S., Rammage, M. and Haddad, T.C. (2019). Implementation of artificial intelligence (AI) for lung cancer clinical trial matching in a tertiary cancer center. *Annals of Oncology*, 30, pp.ii74–ii74. doi:https://doi.org/10.1093/annonc/mdz065.

Mazurek, A., Dziuk, M., Witkowska-Patena, E., Agnieszka Gizewska and Piszczek, S. (2021). 18F-FDG Positron Emission Tomography/Computed Tomography (PET/CT) for Distinguishing Tuberous Sclerosis Complex Lesions from Colon Cancer Metastases. *The American journal of case reports*, 22. doi:https://doi.org/10.12659/ajcr.933320.

Narayan, T., Kumar, S., Kumar, S., Augustine, S., Yadav, B.K. and Malhotra, B.D. (2019). Protein functionalised self assembled monolayer based biosensor for colon cancer detection. *Talanta*, 201, pp.465–473. doi:https://doi.org/10.1016/j.talanta.2019.04.039.

Ramkumar, M.O., Sarah Catharin, S., Ramachandran, V. and Sakthikumar, A. (2021). Cercospora Identification in Spinach Leaves Through Resnet-50 Based Image Processing. *Journal of Physics: Conference Series*, 1717, p.012046. doi:https://doi.org/10.1088/1742-6596/1717/1/012046.

Sakr, A.S., Soliman, N.F., Al-Gaashani, M.S., Pławiak, P., Ateya, A.A. and Hammad, M. (2022). An Efficient Deep Learning Approach for Colon Cancer Detection. *Applied Sciences*, [online] 12(17), p.8450. doi:https://doi.org/10.3390/app12178450.

Shakya, B.R., Shrestha, P., Teppo, H.-R. and Rieppo, L. (2020). The use of Fourier Transform Infrared (FTIR) spectroscopy in skin cancer research: a systematic review. *Applied Spectroscopy Reviews*, pp.1–33. doi:https://doi.org/10.1080/05704928.2020.1791152.

Talukder, Md.A., Islam, Md.M., Uddin, M.A., Akhter, A., Hasan, K.F. and Moni, M.A. (2022). Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning. *Expert Systems with Applications*, 205, p.117695. doi:https://doi.org/10.1016/j.eswa.2022.117695.

Vafapour, Z., Troy, W. and Rashidi, A. (2021). Colon Cancer Detection by Designing and Analytical Evaluation of a Water-Based THz Metamaterial Perfect Absorber. *IEEE Sensors Journal*, 21(17), pp.19307–19313. doi:https://doi.org/10.1109/jsen.2021.3087953.

Yamada, A., Niikura, R., Otani, K., Aoki, T. and Koike, K. (2020). Automatic detection of colorectal neoplasias in wireless colon capsule endoscopic images using a deep convolutional neural network. *Endoscopy*. doi:https://doi.org/10.1055/a-1266-1066.