Predictive Analysis of Colorectal Cancer via CT scans Using Convolutional Neural Networks

Bharathi M P1*, Dr. Samitha Khaiyum2, Dr. Shivakumar Swamy S3

1Research Scholar, VTU-RC, DSCE, Bangalore
2Professor and Head, Department of MCA - VTU, DSCE, Bangalore
3Sr. Consultant Radiologist, Department of Radiology, HCG Cancer Hospital, Bangalore

ABSTRACT: In recent years, the area of Medicine and Healthcare has made significant advances with the assistance of computational technology. New diagnostic techniques including image processing were developed which can help healthcare/clinical experts in many ways. One of the domains is cancer prediction and treatment. Cancer is the world’s second-largest cause of mortality, claiming the lives of one out of every six individuals. Colorectal cancer (CRC) is a common cancer worldwide. It ranks as the third most frequently detected cancer among men and the second among women, with more than 1.4 million new cancer cases every year. This article proposed an improvised CNN approach to analyze the CT scan images of colorectal cancer to predict and classify them into benign(non-cancerous) and malignant(cancerous). In further work the result with tumorous class with patient’s other parameters are integrated for assessing the risk level of cancer. This model results in 96.8% accuracy and minimal error rate. The main contribution of this work is to assist medical fraternity to automatically analyze the CT scan images for prediction and classification of colorectal cancer.

KEYWORDS: Deep Learning, Machine Learning, Neural Network, Convolutional Neural Network, Image Processing, Healthcare, Colorectal Cancer, Heterogeneous Data

1. INTRODUCTION

As per the statistics of American Cancer Society, the colorectal cancer survivors are 65.4%, less than compared to breast cancer (90.35%), prostate cancer (99.6%) [1]. The colon, rectum (together forming the colorectum), and the anus collectively compose the concluding part of the gastrointestinal (GI) tract known as the large intestine. Colorectal cancer refers to cancer disease in large intestine and is also referred as bowel cancer, colon cancer, or rectal cancer, depending on where the cancer cells are present, typically affecting colon and the rectum. Often symptoms are difficult to identify in the early stages. This makes it more important for individuals to learn about the prevention, management, and treatment options.

Nearly 14% of the people suffering from malignant colorectal cancers are assumed to stay alive for five years, when treated at an appropriate time. Still, this is unpredictable because, it depends mainly on few factors like age, gender, the extent to which the cancerous cells have spread to the neighboring organs.[2]

Before delving into an extensive review of current advancements in colorectal cancer diagnosis, it is very important to consider various elements considered for the diagnostic process. These factors include the specific imaging modality employed, the method of diagnostic system utilized, the dataset under examination, and the key metrics employed for evaluating the efficiency of the diagnosis.

Medical images are very important tools to provide better assistance to medical experts. Various imaging techniques are used to find the important features or insights for diagnosis. These techniques are frequently used in combination to obtain sufficient information. Following are the four major screening methods in healthcare.

➢ CT scan: In this current digital era, certain advanced approaches have been developed to deal with colon segmentation. Recently, it is observed that CT colonography uses CT scanning to analyze the interior view of the colon or large intestine [3,4]. It detects the presence of cancer cells that have spread outside of the spectrum.

➢ MRI: To verify if the tumor has been spread through the wall of rectum. It is mainly useful to check whether the tissue left behind after treatment is cancer or not.
➢ **Biopsy (pathology):** The cells or tissues will be removed and viewed under a microscope to identify for signs of cancer. Pathological diagnosing of colorectal cancer requires a pathologist to visually examine digital full-scale whole slide images (WSI). The challenges usually from the difficulty of WSI including image sizes, textures, complex shapes, and histological variations in the nuclear staining.

➢ **PET scanning:** The PET scan involves employing one radioactive substance (tracer) to visualize and detect metabolic activity, showcasing both typical and irregular bodily functions [5]. PET scans aren’t capable of detecting microscopic cells. However, they were used in identifying clusters of tumor cells that may have spread to other organs or tissues within the body.

Among all the above-mentioned imaging modalities CT scans play a vital important role in cancer detection and staging. Numerous studies indicate that the CT imaging characteristics identified by CNNs hold significant predictive potential for oncology-related results [4,6].

### 2. LITERATURE SURVEY

Recent studies have proved that, with the ability of Artificial Intelligence [AI], colon cancer diagnosis can be implemented automatically, with less cost and in a smaller amount time. The AI-based diagnosis is categorized into ML and DL techniques. [7]

Starting from AlexNet in 2012, ZFNet in 2013, VGG-16 in 2014, followed by GoogLeNet in 2014 has proved their efficiency in prediction and classification. 2015 onwards an error rate of less than 4% obtained by the newly proposed algorithms, precisely, an error rate of 3.57% with the ResNet in 2015 and a rate of 3.30% for ResNeXt in 2016. The SENet in 2017 and in 2018 PNASNet-5 have demonstrated even minimum error rates. Over time, CNNs have consistently demonstrated their supremacy in extracting intricate features, both basic and complex, from various architectures. Consequently, they remain the foremost choice for any image classification task in the present landscape especially using image datasets like CT scan or MRI scan.

Convolutional Neural Networks (CNNs) demonstrated their capacity to extract valuable features from medical imaging data, showcasing their recent prowess in contributing meaningful insights to various domains within medical research. [8]. Several studies proved that CNN (Convolutional Neural Networks) can derive CT imaging features with significant predictive potential for outcomes in the oncology field [9].

PCP (Principal component pursuit) [10] used in segmentation of images by applying active contour to background subtracted image. Researchers in [11] proposed a segmentation technique utilizing the contour method introduced a modified active contour framework designed to function independently of edges. This approach provided the distinct advantage of allowing contours to autonomously merge and split.

Researchers utilized a polyp segmentation method rooted in contour region analysis to extract features. They leveraged curvature, polyp color, areas, and edge shapes as characteristics to effectively delineate and identify polyps within the images [12]. Another active contour model used to find pathological tissues using segmented image regions and classified the extracted tissues into three stages of CRC progression using a fully connected CNN [13]. Mainly in CT scan segmentation, Otsu's method can efficiently delineate structures by determining an optimal threshold to separate tissues with varying densities, aiding in the isolation of distinct regions like bones, soft tissues, and organs which helps in medical diagnosis and analysis [14]. During the pandemic it was crucial to diagnose and treat covid-19 patients. It was a challenge to the medical industry to analyze through chest CT images. A study [15] proved the accuracy of three efficient deep learning image classification models with and without Otsu thresholding method with results as illustrated in table below:
### Table 1: Comparison of DL model efficiency

<table>
<thead>
<tr>
<th>Deep Learning Model</th>
<th>Accuracy without Segmentation</th>
<th>Accuracy with Otsu Segmentation</th>
<th>Otsu Segmentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG-16</td>
<td>83.17</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>Resnet-50</td>
<td>70.23</td>
<td>75.08</td>
<td></td>
</tr>
<tr>
<td>MobileNet</td>
<td>72.9</td>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>

### 3. THE PROPOSED METHODS AND MATERIALS

A specialized branch of machine learning is deep Learning, which excels in tackling complex tasks like identifying images, recognizing speech, segmenting images, understanding voice commands, classifying text, and more. Deep learning algorithms replicate the working of a human brain. It takes in data, processes using an activation function, input layer, hidden layers, output layer, loss function, and other important parameters. Image data is fed as input into neurons. This input data is then sent to the next layer with the proper biases and weights. The output is the final value which is estimated from the artificial neuron. In any deep learning model optimization method plays a major role. It discovers the value of parameters which reduce errors while mapping inputs to outputs. Optimizer helps in improving the model’s accuracy and reduces the overall loss. Optimizer helps in improving model accuracy and reduces the overall loss. Selecting a suitable and best optimizer is very difficult because of various parameters machine learning models.

#### 3.1. Convolutional Neural Networks (CNN)

CNNs are widely used in the healthcare industry especially in image analysis [3]. Convolutional neural network (CNN) an approach in deep learning which takes an image as input, it trains the model by learning weights and the biases. The CNN architecture is made up of various layers, i) input layer ii) convolution layer iii) pooling layer iv) fully connected layers. Before we feed images for CNN model, the images must be first preprocessed by resizing, apply normalization, and then augmentation if required.

![Figure 1: Typical Architecture of CNN](image)

### 4. METHODOLOGY

The proposed methodology comprises of 3 different phases: i) Segmentation applied on CT scan images ii) Prediction to identify whether the tumor is present or not iii) Classification of tumor into different stages illustrated in figure (2). Two different CNN models we have used. The first one is improvised CNN used for colorectal cancer prediction. The second one used for the classification of the segmented, preprocessed input images into four different stages of cancer.
A. DATA ACQUISITION

The dataset used for the proposed study are the images of CT scan which are non-contrast axial from 300 patients provided by one of the reputed cancer treatment Centre and hospital in India HCG (Health Care Global Enterprises). Images are of four different classes representing stages of cancer. Approval for the study was received from the Ethics Committee and Academics Committee of HCG along with an MOU signed. All the procedures and discussion with radiologist were performed in full accordance with ethical principles and guidelines of HCG. CT scans images, for which nodule annotations are known. Annotations are the actual nodule locations marked by radiologists. The images in axial view are preferred as it gives the high clarity compared to the other top and side views. The images collected with 220 tumor(malignant) and 80 no cancerous (benign). Additionally, CT images corresponding to colorectal cancer were downloaded from the TCIA imaging Archive using link [16].

B. DATA AUGMENTATION

Efficiency of any Deep Learning models is proportional to the variety in training data, quality, and quantity of dataset. In our case, the dataset size was not sufficient to train an efficient DL model. To expand the dataset size image augmentation technique is been used. Augmentation for an image dataset involves utilizing a generator to create diverse images through transformations like rotation, scaling, flipping, and cropping applied to original dataset. Horizontal flips, vertical flips and random rotations are mainly used to simulate diverse orientations, aiding in enhancing image dataset variability. Augmentation technique is applied to cancerous and noncancerous images separately. Figure (4) shows the visualization of the image augmentation for few sample images from the dataset:

Figure 3: Visualization of Augmentation
In our study, we employed an Image Augmentation Generator with diverse augmentation parameters to enhance our colon image dataset. The generator employed rotation up to 10 degrees, slight shifts in height and width (10% of image size), and shear transformations (0.1). These augmentation techniques aimed to amplify dataset variability for robust colon image analysis.

<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Parameter Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation range</td>
<td>10</td>
</tr>
<tr>
<td>Width shift range</td>
<td>0.1</td>
</tr>
<tr>
<td>Height shift range</td>
<td>0.1</td>
</tr>
<tr>
<td>Shear range</td>
<td>0.1</td>
</tr>
<tr>
<td>Brightness range</td>
<td>(0.3, 1.0)</td>
</tr>
<tr>
<td>Horizontal flip</td>
<td>True</td>
</tr>
<tr>
<td>Vertical flip</td>
<td>True</td>
</tr>
<tr>
<td>Fill mode</td>
<td>nearest</td>
</tr>
</tbody>
</table>

Table 2: Data augmentation parameters.

After augmentation the number of cancerous images is 1500, noncancerous are 450, which we then split into 70% of data for training, 15% of data for validation, 15% of data for testing.

**Stage 1: Segmentation of cancerous region**

Automated colon segmentation has received relatively less attention in research, resulting in fewer advancements or studies dedicated to this specific area within medical imaging. Few approaches have been suggested to support doctors in identifying colorectal cancer. CT scan images of colorectal cancer, varying anatomical positions like colon, intestine, liver rectum demand precise segmentation for focused analysis. Segmentation is pivotal as it enables targeted isolation of tumors across diverse positions. By incorporating segmentation techniques into CNN-based tumor detection systems for cancer in CT scans, the efficiency and accuracy of diagnostic process can be significantly improved. It allows targeted analysis, making the CNN model more effective in identifying and characterizing tumors within complex regions. Whelan and Chowdhury [9] proposed method of using colon geometrical features for automatic colon segmentation from CT data [17].

Image segmentation stands as a pivotal component in the realm of medical image processing. It encompasses a methodical approach to extract the ROI (Region of Interest) within medical images through either automatic or semi-automatic procedures. Across medical applications, various image segmentation techniques have been deployed extensively to delineate tissues and other body organs. Certain necessary steps are followed in the proposed segmentation technique.

**Gaussian denoising**

Images captured from various sources or sensors often contain noise due to conflict in the acquisition process. Image segmentation algorithms can be sensitive to noise. Noisy images might have abrupt pixel intensity changes that can lead to erroneous segmentation results. Several techniques are already in use to normalize the images. In our study we used Gaussian denoising which reduces the noise, improves the segmentation quality, preserves edge and ensures better feature extraction. A study [18] proposed a model with increased segmentation accuracy using Gaussian filter for CT and MRI image denoising. They could successfully preserve the edges and other structures of image effectively. The cv2.GaussianBlur() function is used to apply a Gaussian blur to the grayscale image . The sigma X=0.5 parameter sets the standard deviation in the X direction of the Gaussian kernel.

**Segmentation using Otsu's thresholding method.**

Thresholding, a prominent segmentation technique, relies on the premise that images encompass distinct regions with varying grey levels and histograms are applied to observe peaks and valleys. Employing a threshold value within the histogram partitions pixel intensities into two categories: the
"foreground" comprises pixels with intensities equal to or higher than the threshold, while the "background" comprises pixels with intensities lower than the threshold, enabling effective segmentation [19]. A study demonstrated on the performance of Otsu method on large set of noisy images. [20]

Segmentation approach partitions the images directly into regions based on intensity values or/and other properties. Thresholding creates all binary images which are from grey level by converting entire pixel values below some value of threshold to zero. All other pixels above that threshold value to one. Assume \( g(x, y) \) which is a threshold version of \( f(x, y) \) at a value global threshold \( T \) and is defined as:

\[
g(x, y) = \begin{cases} 
1 & \text{if } f(x, y) \geq T \\
0 & \text{otherwise}
\end{cases}
\]

From pixel 0 to 255 we can find the optimal threshold value by evaluating and calculating between-class variance. For the same formula used is:

\[
\text{var}(T) = P_0(T) \times P_1(T) \times (m_0(T) - m_1(T))^2
\]

Where \( P_0(T) \) and \( P_1(T) \) represent probability values of the background and foreground regions, respectively. Also, we use \( m_0(T) \) and \( m_1(T) \) as the mean grayscale intensity values of the background and foreground regions, respectively. This creates a color mask highlighting a particular ROI (Region of Interest). Additionally, we apply connected component analysis to identify distinct components and isolate the largest component assumed to represent a tumor or critical area after. It is necessary to identify the borders and other edges after applying Otsu technique for which few simple steps to be followed [21]. Further processing involves extracting the white regions from this identified area in HSV color space. cv2.bitwise_and () operation combines this mask with original HSV image.

![Figure 4](image.png)

Figure 4: Segmentation of cancerous(malignant) and non-cancerous(benign) region

The suitable features are identified and extracted from preprocessed endoscopic images representing the benign and malignant colorectal polyps containing CT scans with normal and cancerous, for which nodule annotations are known. Annotations are the actual nodule locations marked by radiologists. The countered images are been used to analyze the region of malignant cells in the region of colon or/and rectum.

Benign or cancerous tumors do not spread to different organs of the body, whereas non-cancerous or malignant tumors spread to different organs by affecting nearby healthy tissues. To identify the type of tumor the segmented image is used as input in this stage. The improvised CNN model is used for prediction which has 4 convolutional layers, 3 dense layers. This model produced 96.8% accuracy.

The first and last two layers are modified to tune the model. To match the size limitations of the proposed model, images were resized to 224×224×3.
Table 3: Architecture of proposed Improvised CNN

<table>
<thead>
<tr>
<th>Layer Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input Layer</strong></td>
<td>Images of size 224x224 pixels with 3 color channels (RGB)</td>
</tr>
<tr>
<td><strong>Convolutional</strong></td>
<td></td>
</tr>
<tr>
<td>Layers + RELU</td>
<td>• Conv2D with 64 filters, each using a 3x3 kernel + ReLU activation.</td>
</tr>
<tr>
<td></td>
<td>• Conv2D with 64 filters, each using a 3x3 kernel + ReLU activation.</td>
</tr>
<tr>
<td></td>
<td>• Conv2D with 128 filters, each using a 3x3 kernel + ReLU activation.</td>
</tr>
<tr>
<td></td>
<td>• Conv2D with 128 filters, each using a 3x3 kernel + ReLU activation.</td>
</tr>
<tr>
<td><strong>Pooling Layers</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• MaxPooling2D layer after the initial two convolutional layers with a pool size of (2, 2).</td>
</tr>
<tr>
<td></td>
<td>• MaxPooling2D layer after the next two convolutional layers with a pool size of (2, 2).</td>
</tr>
<tr>
<td></td>
<td>• MaxPooling2D layer after the last convolutional layer with a pool size of (2, 2).</td>
</tr>
<tr>
<td><strong>Flatten Layer</strong></td>
<td>Flatten layer to transform the 2D feature maps into a 1D vector</td>
</tr>
<tr>
<td><strong>Dense Layers</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Dense layer with 64 neurons + ReLU activation.</td>
</tr>
<tr>
<td></td>
<td>• Batch Normalization layer to normalize the activation of the previous layer.</td>
</tr>
<tr>
<td></td>
<td>• Dropout layer with 0.1 dropout rate to prevent overfitting.</td>
</tr>
<tr>
<td></td>
<td>• Output layer is with a single neuron and a sigmoid activation function for binary value</td>
</tr>
<tr>
<td><strong>Loss Function</strong></td>
<td>Binary Cross entropy</td>
</tr>
<tr>
<td><strong>Optimizer</strong></td>
<td>Adam</td>
</tr>
</tbody>
</table>

The first convolutional layer is the input layer that applies 64 filters to the inputted image of size 224x224 pixels with 3 color channels (RGB). It uses relu activation function to introduce non-linearity. Another convolutional layer with 64 filters extracts features from previous layer’s output. The third layer with 128 filters applies the activation function ReLU to capture more complex patterns. After each convolutional layer, max-pooling is applied with 2x2 window, and a stride value of 2 helps in reducing the spatial dimensions of data and used to retain the important features.

![Figure 4: Sample of Pooling operation in max-pooling Layer](image)

Flatten layer flattens the output/result from previous layer into a 1D array, preparing it to be fed into a densely connected layer. It produces 4135436 parameters. During training Batch Normalization() normalizes the activation of previous layer for faster convergence. A regularization technique where 10% of the neurons from the previous layer are randomly dropped during training to avoid overfitting. The output layer (final layer) with a single neuron and a activation function sigmoid outputs a value ‘0’ or ‘1’ indicating Benign (non-cancerous) and Malignant respectively.

**Experimental setup**

This experiment was carried out on an Intel i7 processor with 16GB memory support controlled by the Windows operating system. A graphic processor having 896 CUDA cores, a base clock speed of 1395 MHz, and a dedicated memory of 4GB GDDR6 was used. A few important libraries that were used for the study are keras, tensorflow, watershed etc.
5. RESULTS & DISCUSSIONS

The dataset collected from HCG hospital without any patient identity. CT scans of colorectal cancer encompass the abdomen, liver, pelvis, focusing on the colon (large intestine) and rectum are utilized in the study for predictive purposes. The axial CT scans corresponding to normal, benign polyp, adenomatous polyp and moderately and highly malignant polyp conditions were considered for the study. The radiologist from HCG made the annotations and specified images with the outlined contours of the tumors. These images helped us to validate the result as well. The countered images are then used for further study.

The segmentation technique used is simple, efficient, and resulted in a notable reduction in loss compared to the models proposed and used by various researchers. According to the study [20] Otsu method is used effectively to find the threshold value automatically for the large set of CT images. In our study segmentation technique, which plays a major role in increasing the model performance.

Initially a basic CNN model used with 3 convolutional layers which obtained 92% accuracy. Then improvised CNN with four convolutional layers resulted with 96.8% accuracy. The proposed improvised CNN’s performance of the has been measured using the most essential parameter accuracy. Also, we use confusion matrix to translate the binary classification corresponding to True Positive (TP), True Negative (TN), False Negative (FN), False Positive (FP), F1-score as defined below:

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}
\]

\[
F1\text{-score} = \frac{TP}{(TP + 1/2(FP + FN))}
\]

![Confusion Matrix](image)

Figure 5: Confusion matrix for the test dataset

The efficiency of any model always depends on the size of the dataset. In a study a CNN model used for prediction of colon cancer produced 95.8% accuracy for 296 images.[22] Whereas our improvised CNN produced 96.8% accuracy for 1500 images and for 26 epochs with F1 score of 96%.

![Accuracy and Loss](image)

Figure 6: Accuracy and Loss
Ethics Approval

The CT images collected and the discussion with respect to specific cases with radiologist from HCG (Health Care Global Enterprises) study were provided in accordance with acceptance and approval of the ethics committee.

CONCLUSION AND FUTURE WORK

The proposed Improvised CNN model demonstrates promising potential in accurately predicting colorectal cancer using axial CT images collected from HCG hospital. To enhance the image volume, we have downloaded images from TCIA (Cancer Imaging Archive). We have used different pre-processing steps like cleaning, frames extraction, noise removal, handling of imbalanced categories, image enhancement, cropping and resizing the images suitable for the requirement. Compared to several existing prediction models our model resulted in increased accuracy of 96.8%.

As a future work, these predicted cancerous images are used to classify according to cancer stages. This model can be efficiently used by medical experts.

Acknowledgments

We gratefully acknowledge HCG (Healthcare Global Enterprises Ltd) for providing the dataset with necessary support and guidance. Also thankful to Dr. Shivakumar Swamy S, Sr. Consultant Radiologist, Department of Radiology, HCG Cancer Hospital, Bangalore for the valuable inputs and guidance.

REFERENCES


DOI: https://doi.org/10.15379/ijmst.v10i3.3365

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.