IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

Detection of Colon Cancer using Deep Learning Techniques

Prof. Neetha Das[1], Akshay K H[2], Amisha P[3], Arjun Ganesh G[4], Deepak V[5]
Students, Department of Computer Science and Engineering[1,2,3,4]
Assistant Professor, Department of Computer Science and Engineering[5]
Bangalore Institute of Technology, Bengaluru, India

Abstract: *Cancer is the most common disease in the world. Cancer is a disease in which some somatic cells grow out of control and spread to other parts of the body. Among the types of cancer, colon cancer is the third most frequent cancer in the world. Polyps which are developed inside the human body parts are the most common cancer cause. These polyps are initially benign, they later grow to become cancerous over time, thus early detection and screening are critical for the patients. Artificial Intelligence is being used to diagnose cancer and deep learning techniques are used for prediction. The motive of the paper is to propose a model to implement a deep learning model like Convolution Neural Network (CNN), and InceptionV2 Network to predict cancer and non-cancerous. The model helps in the classification framework and prediction of two types of colon cancer: they are colon adenocarcinoma and colon benign cancer. The model is deployed for the histopathological image data set, along with the deep learning model*

Keywords: Insecption V2 Network; colon cancer; colon adenocarcinoma and colon benign; machine learning, and CNN

I. INTRODUCTION

In the present world, there are many diseases affecting living beings. Cancer is one of the diseases which is causing major concern in the medical field. Cancer is a term used to describe a group characterized by abnormal cells growing uncontrollably. There are many types of cancer, which are classified based on the type of tissue where cancer originated. Colon cancer is also a type of cancer, which affects the colon (the last part of the large intestine). Colon cancer is the third most common type of colon cancer. The colon or a large intestine is where the body draws the water and minerals from solid waste. The waste subsequently exists in the human body via the anus after passing through the rectum. Colon cancer is a common digestive system cancer that arises in the colon most commonly at the junction of the rectum and sigmoid colon. Colon cancer is mainly developed from the polyps that start originating in the tissue of the colon, these polys have a high chance of developing into cancer. The Polyps develop into cancer and can spread to other organs if they are not discovered and diagnosed on time. Polyps can be found via screening tests and removed before they turn into malignancy. Colorectal cancer screening also aids in the early detection of the disease, when treatment is required. Colon polyps are examined using Colonoscopy techniques to capture the diseased part in the body.

The idea of the project is to build a model using Inception Networks, model for the detection and classification of two types of cancer: colon adenocarcinoma and colon benign.

- Colon Adenocarcinoma: Tumors that begin inside the internal organs of a human being are called adenocarcinomas. The word “adeno” comes from the Greek word which means gland. These cancers begin in glandular cells. They can develop in a variety of organs, including the lungs and the breast. Early tumors in colorectal cancer begin as tiny adenomatous polyps that expand and eventually transform into malignant tumors. Adenocarcinomas account for the great majority of colorectal cancers.
- Colon Benign: Colon polyps that are frequently common are benign lesions. These tumors begin in the muscular tissue of the intestine, they are initially benign but eventually develop into cancer.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

The Deep Learning models like Convolution Neural Networks (CNN) help in predicting the presence of cancer using the tissue image of the colon. For image segmentation and classification, CNN and its variants have been well known for handling complex data. In the field of diagnosis, Artificial Intelligence (AI) has shown tremendous promise and provided us with a viable alternative to traditional diagnostic approaches. In the applications that handle the images, such as Image Classification information (Image Net), computer vision, and Natural language processing (NLP) produces the results are truly remarkable.

II. PROPOSED SYSTEM

This research aims to use the colon tissue image to predict the existence of cancer using the Convolutional Neural Network (CNN) approach. CNN is widely used for image recognition. It has been known to handle complex data for image classification. Feature extraction plays a major role in segmenting objects. Many polyp segmentation methods proposed in the literature consider texture, color, and shape features. All these features are considered low-level features and not sufficient for precise segmentation of polyp regions. Convolutional neural network (CNN) architectures are proven to be efficient in the automated extraction of features for various computer vision applications. Early diagnosis of colorectal cancer is achievable. Colonoscopy is the primary method for screening and preventing polyps from becoming cancerous. However, identifying the disease at an early stage significantly improves the chances of survival. A Cancer diagnosis can be automated by using the potential of Artificial Intelligence (AI), which allows us to assess more cases in less time and cost. Implementation of this model will help medical professionals to develop an automatic and reliable system capable of identifying various types of colon cancers. If polyps are not detected and removed, they can turn out to be cancer and can spread into other organs. Chances of survival can be increased if the polyps are detected and removed when they are very small. Therefore, it is advised by many doctors have colonoscopy regularly. The system proposes a model for identifying and classifying colon cancer using CNN.

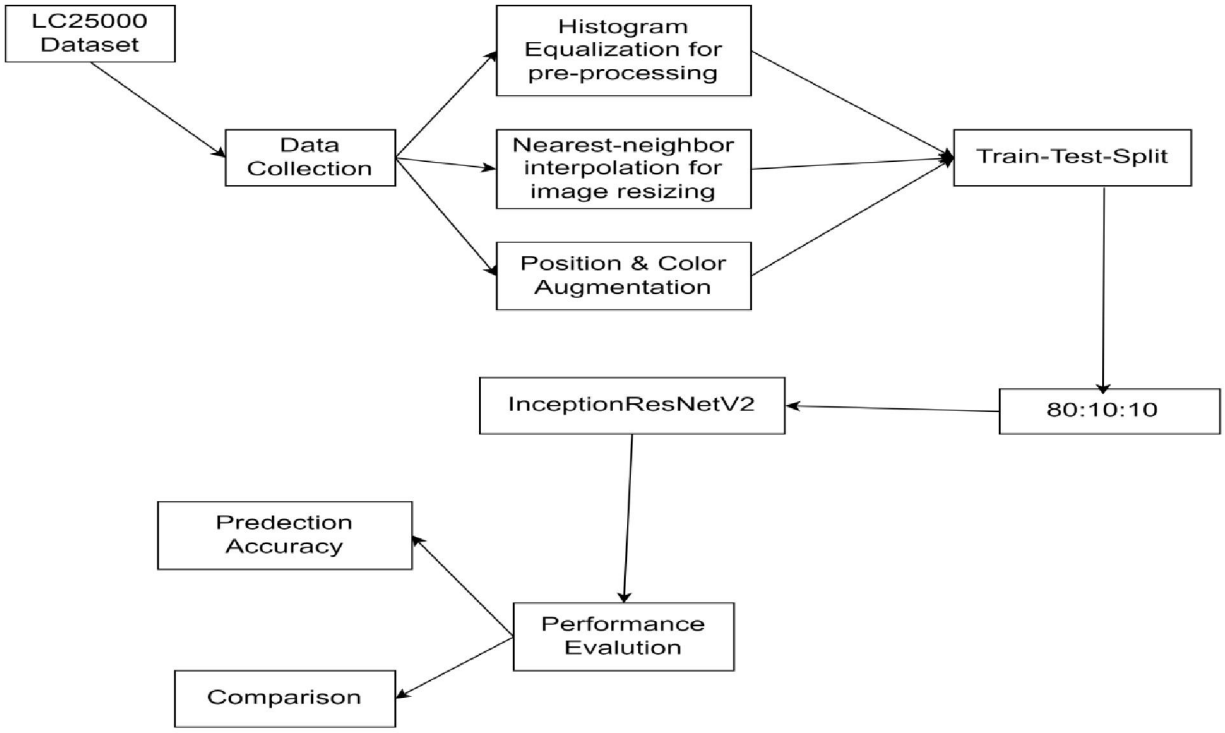

Fig1:System Architecture

2.1 The Colon Cancer Datasets

The dataset collection is fundamental to developing a deep learning model. The image dataset used for colon cancer is Histopathological Images. The dataset collection and how it is preprocessed are shown in Fig 4.2. It is acquired from the LC25000 dataset, which is a histopathological imaging dataset for lung and colon cancer. The data collection

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

compiled by Andrew A. Borkowski and colleagues has 25,000 colour images of five different types of both lung and colon tissues. The downloaded dataset of histopathological images included variants of both cancer i.e., lung and colon image datasets classified into five subclasses three for the lung cancer variants and two for colon cancer variants.

Original image dimension were 1024 x 768 pixels, they were resized to 768 x 768 pixels to have a square dimension before augmentation procedures.10000 images of colon tissues (5000 images of colon adenocarcinoma and colon benign cancer class each).The data collected after data augmentation from the histopathological dataset is split into train and test data. The data is first split into train and test data in a ratio of 80:20, in that test ratio 10:10 is divided for test and validation data. These images were created from a sample of 500 HIPAA validated photographs of colon tissue. All images in the dataset have a resolution of 768x768 pixels which are saved in jpeg image format.

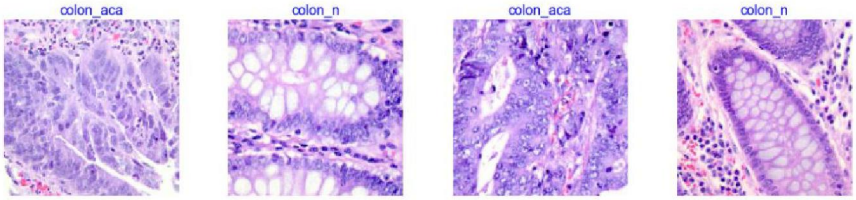

Fig 2: Dataset

Type of Cancer	Class name	Number of samples
Colon adenocarcinoma	col_aca	5000
Colon benign tissue	col_n	5000

III. INCEPTION RESNETV2 NETWORK

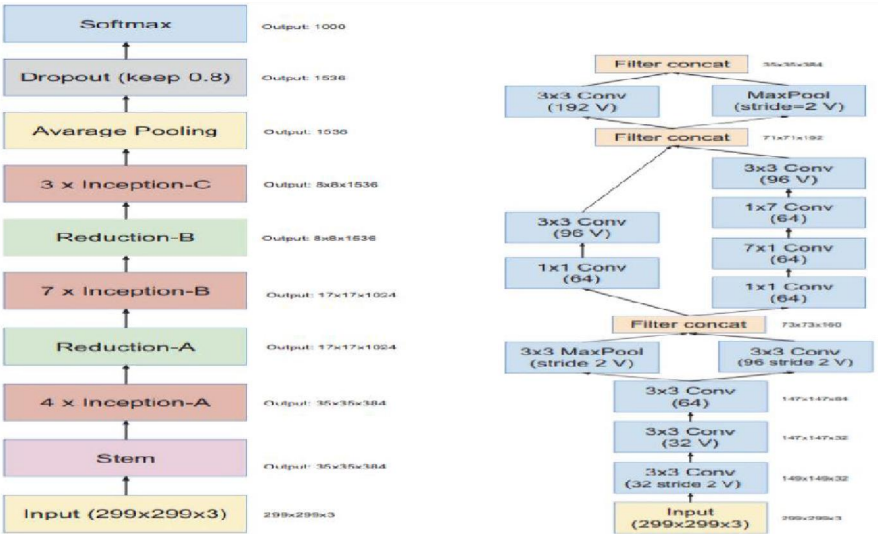

Fig 3: Inception-ResNet-V2

Inception-ResNet-v2 is a type of deep convolutional neural network (CNN) model. The model is to identify and diagnose colon cancer from the histopathological image dataset. The model is trained for the histopathological dataset of colon adenocarcinoma and colon benign tissue images. InceptionResNetV2 is trained using batch normalization, which improves the speed of the training. The model decreases the filter size using factorization, which helps to reduce the chances of overfitting problems. The model consists of block, stem, and Inception-ResNetV2 with Softmax activation, average pooling, dropout, and a dense layer. This substantially lowered computational complexity by replacing kernels with numerous smaller convolutional kernels. As the number of convolutional layers increases and the network depth enhanced performance accuracy.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

IV. PARAMETERS, RESULTS AND ANALYSIS

In this section, we have reported the results of our model and shown the analysis based on parameters such as model accuracy, precision, recall confusion matrix.

Table 1 shows comparative study of different classifier with respect to accuracy, precision and recall as parameter.

Classification Report:

	precision	recall	f1-score	support
colon_aca	1.00	0.99	0.99	240
colon_n	0.99	1.00	0.99	260
accuracy			0.99	500
macro avg	0.99	0.99	0.99	500
weighted avg	0.99	0.99	0.99	500

Confusion Matrix is one of the tools for evaluating the behaviour a binary classifier. For better visualization of results, we have used heat maps for the model.

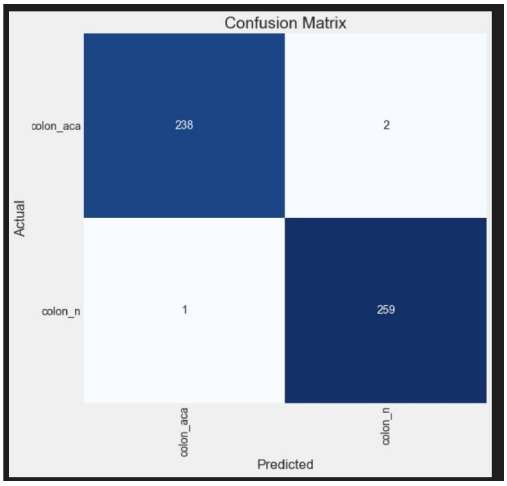

Fig 4: Confusion Matrix

Fig 5:Training and Validation Loss and Validation Accuracy

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

V. CONCLUSION

The Data was collected from the histopathological images. The dataset is used to train with the InceptionResNetV2 network with hyperparameters for achieving better accuracy. With the mentioned deep learning network, the model achieved a validation accuracy of 99.6%.The Early-stopping feature with patience helped to monitor the performance and to stop the epoch if it's causing an overfitting problem. An overall model can provide a better accuracy as par compared to theprevious models and methods deployed. The prediction model can be used in healthcare applications. The model can predict colon cancer and can also classify it as colon adenocarcinoma or colon benign.

REFERENCES

[1] Xiaoyong Yang, Qianxing Wei, Changhe Zhang, Kaibo Zhou, Li Kong and Weiwei Jiang "Colon Polyp Detection and Segmentation based on improved MRCNN" IEEE Transactions on Instrumentation and Measurement volume: 70,2021.

[2] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), pp. 2278-2324.

[3] DevviSarwinda; AlhadiBustamam; Radifa H. Paradisa; Terry Argyadiva; WibowoMangunwardoyo "Analysis of Deep Feature Extraction for Colorectal Cancer Detection" 4th International Conference on Informatics and Computational Sciences (ICICoS),2020.

[4] R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "GradCAM: Visual Explanations from Deep Networks via Gradient-Based Localization," 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618-626, doi:10.1109/ICCV.2017.74.AIDS Res. Ther., vol. 17, no. 1, pp. 1–13, 2020.

[5] F. Bray and I. Soerjomataram, "The changing global burden of cancer: transitions in human development and implications for cancer prevention and control," in Cancer: disease control priorities, vol. 3, Washington: World Bank Publications, 2015, pp. 23-44.

[6] S. Rathore, M. Hussain, A. Ali, A. Khan, A recent survey on colon cancer detection techniques, IEEE/ACM Trans.Comput. Biol. Bioinform. 10 (2013) 545563.

[7] E. Ozdemir, C. Sokmensuer, and C.G. Demir, A Resampling- Based Markovian Model for Automated Colon Cancer Diagnosis, IEEE Trans. Biomedical Eng., vol. 59, no. 1, pp. 281-289, Jan. 2012.

[8] A.C. Simsek et al., Multilevel Segmentation of Histopathological Images Using Cooccurrence of Tissue Object, IEEE Trans. Biomedical Eng., vol. 59, no. 6, pp. 1681-1690, June 2012.

[9] E. Ozdemir and C.G. Demir, A Hybrid Classification Model for Digital Pathology Using Structural and Statistical Pattern Recognition, IEEE Trans. Medical Imaging, vol. 32, no. 2, pp. 474-483, Feb. 2013.