DEEPLEARNINGFORCOLORECTALCANCER DIAGNOSIS: A CLASSIFICATION APPROACH

[1]KHADAKKAR POONAM RAVINDRA [2]DR.RAHUL M. PATIL
[3]DR.VINODKUMAR P. PATIL

Department of Electronics and Telecommunication Engineering,
NES'S Ganagamai College Of Engineering, Nagaon
poonamnikume39@gmail.com

ABSTRACT:

Cancer refers to a wide range of diseases marked by the unchecked growth and proliferation of irregular cells within the body. If not addressed, cancer can disrupt essential bodily functions and pose a serious risk to life. Colorectal cancer is one type ofcancer. Colorectalcancer is one of the mostcommonandfatal malignant neoplasms worldwide.Traditional diagnostic methodsare basedon histopathological analysis, whichconsumes timeand affectsinter-observervariation. Deeplearning, inparticular, convolutionalneural networks (CNN$_S$), hasemergedasapowerfultoolforautomatingandenhancingCRCclassification.

Conventional diagnostic techniques depend on histopathological assessment, which is time-consuming and susceptible to differences in interpretation among observers. Deep learning, especially convolutional neural networks (CNNs), has become an effective method for automating and improving the classification of colorectal cancer. This paper evaluates different deep learning strategies utilized for the diagnosis of colorectal cancer. ThedifferentarchitecturesincludesGoogLeNet,AlexNet,VGG,Inception,ResNet.Outofmostofthetechniques availableResNetprovestobebetteramongstall.

***KEYWORDS**:Colorectalcancer,machinelearning,deeplearning,CNN,Alexnet,GoogleNet,VGG, Inception, ResNet.*

1. Introduction

The term"cancer" is used to refer to a broad category of diseases that can affect different body organs. These cells have the potential to develop into cancer as they divide and spread throughout the organs, which could result in death. Cancer ranks as one of the top causes of death, just behind cardiovasculardiseases.Worldwide,colorectalcancer remains the third most commonly studied type of cancer and the third leading cause of deaths linked to cancer. The development of personalized medicine is a trending issue in cancer treatment, as it has the potential to enhance the effectiveness of therapy by tailoring it to the unique characteristics of each patient's cancer[1].

According to a National Cancer Institute (NIH) figure, there will be about 606,520 cancer-related deaths and 1,806,950 million new cases in the United States in 2020. The World HealthOrganization estimates that by 2040, 28.4 millionnew cases of cancer will be diagnosed worldwide, making it the leading cause of death worldwide[2].

Roughly 70% of cancer fatalities occur in low and middle-income nations. Merely 26% of nations with low income possess the necessary pathological materials required for cancer diagnosis accessible to the general public, based on data from 2016; affluent countriesmightprovidecancerscreeningand assessment to more than ninety percent of their populace[3-5].

Either the colon or the rectum is where colorectal cancer begins. Depending on where they begin, these cancers may also be referred to as rectal or colon cancer. Due to their many similarities, colon cancer and rectal cancer are frequently grouped together.

Colorectal cancer, commonly referred to as large intestine cancer (CRC) is the fourth leading cause of morbidity and the third primary cancer deaths globally[6].

Deep learning (DL), a branch of machine learning (ML), is frequently utilized in researchfields, including medical image analysis, computer vision, speech recognition, and natural language understanding. Over the last six years, deep learning has garnered significant interest because of enhanced computingcapabilities,decreasedhardwareexpenses, andthe developmentofnumerousnewdatasets.Deep learning algorithms excel in identifying and diagnosing cancers, in addition to tumor segmentation, as they can directly extract advanced features from unprocessed images. Deep learning techniques can assist doctors by providing alternative insights and emphasizing regions associated with images [7].

2. Background

Cancer

One of the most critical illnesses is cancer. Human bodies are formed of cells.Cells are replaced by other freshly formed cells continuously. The cells are divided by cell division in which one cell is divided into smaller cells.All theinformation abouta cellisstoredinDNAinthe formof genes. Asthe cell divisiontakesplace,theDNAandtheinformationare passed to the newly formed cell. Mutation refers to mistakes that take place while passing the DNA further. Mutations are likely to happen while passing the information. The information stored in DNA instructs the cells about the time when it should stop the cell division.But the cell division does not stop if the cells are mutated, which causes the abnormal growth of cells.This abnormal growth of cells is called cancer.

Thistypeofabnormalgrowthofcellscantake place in any part of the body i.e, lungs, skin, brain, breast, colon, rectum, etc.When cells divide uncontrollably and form a cluster, a tumor results. However, since it is a benign tumor, treatment may not be necessary. A malignant tumor (cancer), on the other hand, is a type of tumor that spreads to other parts of the body and can have a variety of effects on your health [8]. Although there isn't a single way to avoid cancer, there are several variables that can lower your risk, including stopping smoking, getting vaccinated, getting regular checkups, keeping a healthy weight, exercising frequently, eating ahealthy diet, and getting diagnosed early[9].

ColorectalCancer

Colorectal cancer refers to the abnormal growth of cellstakingplaceinthecolonorrectum.Thecolonor rectum are efficacious parts of the large intestine.This abnormal growth of cells in the large intestine results in inefficient working of the large intestine,i.e, it leads to problems related to the digestive system. Most of the functioning of body parts relies on the digestive system as it results in the extraction of nutrients. So, to safeguard the proper functioningof the digestive system, it is essential to detect colorectal cancer at an early stage[10,11].

ColorectalCancer

Deep learning (DL), a branch of machine learning (ML),isadevicefrequentlyutilizedinresearchfields like healthcare image manipulation, machine vision, voice analysis, and processing of natural language. Over the last six years, profound. Learning has garnered significant interest because of the rise in processing capability alongside decreased hardware expenses and the development of numerous new datasets.Becausedeeplearningalgorithmscan directly extract high-level features from raw images, they are effective in tumor segmentation and cancer detection and diagnosis. By highlighting regions associated with images and providing secondary ideas, deep learning techniques can assist doctors. Additionally, among medical techniques, it has even been demonstrated that a single deep learning model is useful for diagnosis[7].

3. On-handmethodologies

A convolutional neural network (CNN) is a form of deep learning method. Deep learning is a machine learning technique that employs artificial neural networks to learn from data. Deep learning can address numerous challenges, such as image identification, natural language understanding, and voice recognition. Deep learning models are capable of identifying data patterns, such as intricate images, written text, and audio. The names of the most well-known and widely used CNN architecture types in deep learning are AlexNet, GoogleNet, ResNet, LeNet,VGGNet,MobileNets,ZFNets,Inception,etc.

4. LiteratureSurvey

For the classification of colorectal polyps, multiple studies have been proposed, which are automated methods. For adenoma grading, two levels of classification are given, which are i)patch level classification and ii)whole slide image(WSI) level classification. But because of the high quality of histopathological images, inputting of whole slideinto the classification model is impractical. However, downsizing WSI into smaller images can result in the loss of comprehensive information, which may be crucial for processing. Hence, the division of WSIinto small patches and then performing two-level classifications is the most optimized way to solve the mentioned issue[12].

A study that looked into the use of DL algorithms for CRC detection was published in the literature. Sakr et al.[13], for example, provide an effective DL-based. The suggested technique extracts features from colon histology images using a deep convolutional neural network (CNN). A fully connected neural network is then fed the retrieved features to classify them. The research employed a dataset containing 1000 pathological images of colon cancer, which underwent preprocessing to eliminate background noise and improve image clarity. Followingpreprocessing,theimageswereaugmented to expand the dataset and enhance the generalization capabilities of the model. A transfer learning method was applied to extract features from the preprocessed and augmented images, utilizing a pre-trained VGG-16 model.

Korbar et al. [14] created several deep-learning models to categorize five common types of polyps—hyperplastic,sessileserrated,traditional serrated,tubular,andtubulovillous/villous—inlight of the recent advancements in deep-learning techniques.Usingadatasetof2074patches,they evaluatedtheperformanceof AlexNet,VGG, GoogLeNet, and a variant of ResNets. According to their tests, ResNetD fared better than the others and offered patch-level classification accuracy of 91.3%. Additionally, they evaluated their approach on 239 separate whole-slide H&E-stained images and obtained 93.0%overallperformance accuracy, 89.7% precision, 88.3% recall, and 88.8% F1 score.

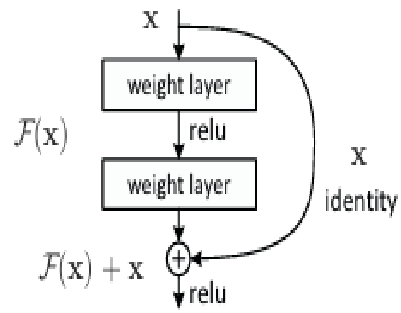
Nucleidetectionandclassificationarethe differenthistologicalimageanalysistasksthatare essential to retrain the pre-trained model. ConvolutionalNeuralNetworks(CNNs)arewidely usedtoachievethistargetofpre-trainedmodel.They employedtwodistinctapproachestoretrainthepre-trainedmodeltoarriveatthefinalresults,whichare theK-foldcross-validationmethodandthenormal classificationmethod.Thismodel canclassifythe datasetinavarietyofways.TodescribetheNormal ClassificationMethod(NCM),thedatasetwillbe splitintotwosections,referredtoastrainingdataand validationdata.So70%ofdatafortrainingand30% forvalidationischosen.Consequently,thismodel achieved excellent performance as it will be educated on the highest volume of information that may result in equipping a neural network with a robust ability to categorize colon cancer[15].

Additionally, Christian Szegedy[16] employedasomewhatbetter-quality,deeper,and widerInceptionnetwork;yet,its inclusioninthe ensembleappearedtoonlysignificantlyenhancethe outcomes.Theyleaveoutthespecificsofthat networksinceempiricaldataindicatesthattheprecise architectural characteristics have a negligible impact. Six ofthe seven models in our ensemble were trained using this network, which was trained using various image patch sampling techniques. Rectified linear activation is used in all convolutions, including those found in the Inception module . Our network's receptivefieldhasazeromeanandmeasures224x 224 in the RGB color space. The numbers "#3×3 reduce" and "#5×5 reduce" indicate how many 1×1 filters were employed in the reduction layer beforethe 3×3 and 5×5 convolutions. Following the built-in max-pooling, the number of 1x1 filters in the projection layer is visible in the pool projection column. Rectified linear activation is also used by all of these reduction/projection layers.

DifferentCNNarchitecturesareexplainedby PrashantNarayanSholapur[17] fortheeffective detectionofcoloncancer.Outofmanyavailable methodologies,hefocusedonInceptionV3model. Heretrievedinformationfrom[18],whosaidan artificial neural network with multiple layers is the Convolutional Neural Network modelInceptionV3.To determine whether colon cancer is present, the model is integrated with the collection of histologicalimages.Convolutionallayers,average andmaximumpooling,concatenationlayers,fully linkedlayers,dropoutlayers,andSoftMaxactivation makeuptheInception3modelinourensemble. Constructingthemodelisstraightforward,withthe installation performed automatically. The alterationof weights will occur when the model is built. For weights, the parameter is allocated. The ImageNet, which will enable the use of the weights that are pre-trained on ImageNet.

5. ProposedMethodology

Therearevariousarchitecturesavailablein deep learningtechniques. Inthispaper,the classificationofcolorectalcancerisproposedusing ResNetarchitecture.ResNet standsforResidual Network.Aphenomenonknownasthe"vanishing gradientproblem"happens whendeepneural networksarebeingtrained. Thegradientsthatare usedtoupdatethenetworkgetsolittleor"vanish"as theyarebackpropagatedfrom theoutputlayers tothe previous layers. To overcome this problem, a deep learning residual framework is used. ResNet specificallyallowsafewstackedlayerstofita residualmappingratherthan relyingonthemto directly match a specified underlying mapping. With shortcut connections, feedforward neural networks can implement the formulation of F(x)+x[19]. Connections that skip one or more layers are known asshortcutconnections[20,21].Inthisinstance,the shortcutconnectionsmerely carryoutidentity mapping,addingtheiroutputstothestackedlayers' outputs.



Feedforward neural networks with "shortcut connections" can be used to achieve the formulation ofF(x)+x(Figure1).Theoutputsoftheshortcut connectionsaresimplyaddedtotheoutputsofthe stackedlayers,andtheycarryoutidentitymapping (Figure1).Identityshortcut linksdon'tincrease computational complexity or add unnecessary parameters.

6. Conclusion

The proposed model of classification of Colorectal cancer using deep learning techniques proposes the early detection of colorectal cancer and can be implemented best using Residual Network, i.e ResNet. As ResNet can overcome the vanishing or disappearing gradient problem that can occur in any CNN architecture. Any biological orhistopathological images i.e, any image related to medical imaging, are mostly asymmetric. So, for the processing of medical images, a very deep neural network is necessary. ResNet can learn more complex features because it can train very deep networks, which is the most important advantage of ResNet when it comes to medical images. ResNet's quality of reducing the vanishing gradient problem made it a winner in the ImageNet Large Scale Visual Recognition Challenge(ILSVRC). The learning rateofResNetiswithintheidealrangeoflearningrate, i.e0.001 to0.01.Hence,itisoneoftheadvantagesof ResNet over other architectures that makes it better than other architectures. It is also observed that ResNet resulted in 93% accuracy for patch-level classification and about 93% accuracy in whole slide images, which is again a better result. Along with the accuracy, it also provides good precision, recall, and F1 score, which are significant parameters in Deep learning. The model's accuracy can be extended to 95%,whichisanoutstandingoutcome.Hence,taking into consideration the overall study, I would like to summarize that ResNet architecture can give better resultsfortheclassificationofcolorectalcancerusing deep learning techniques.

6. References

[1] E.N. Srivani and G. Seshikala " Colon Cancer Detection using Deep Learning ". 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)

[2] Algorithm" 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., & van der Laak, J. A. W. M. (2017). A survey on deeplearninginmedicalimageanalysis.*Medical Image Analysis, 42,* 60-88. https://doi.org/10.1016/j.media.2017.07.005.

[3] S. Mehmood et al., "Malignancy Detection in Lung and Colon Histopathology Images Using Transfer Learning With Class Selective Image Processing,"inIEEEAccess,vol.10,pp.25657-25668, 2022.

[4] Y. Li, F. Zhang and C. Xing, "Screening of Pathogenic Genes for Colorectal Cancer and DeepLearningintheDiagnosisofColorectal Cancer,"inIEEEAccess,vol.8,pp.114916-114929, 2020.

[5] C.T.SariandC.Gunduz-Demir,"Unsupervised Feature Extraction via Deep Learning for HistopathologicalClassificationofColonTissue Images," in IEEE Transactions on Medical Imaging, vol. 38, no. 5, pp.1139-1149, May 2019.

[6] P. Rawla, T. Sunkara, and A. Barsouk, "Epidemiologyofcolorectalcancer:incidence, mortality, survival, and risk factors," Gastroenterol.Rev.,vol.14,no.2,pp.89–103, 2019,doi:10.5114/pg.2018.81072.

[7] Allan C Muzenda, Ottilia Faith Chimpeni and Fine Masimba "Colorectal Cancer Prediction Using Deep Learning" 2023 2nd Zimbabwe ConferenceonInformationCommunicationand Technology (ZCICT).

[8] Pancreatica.WhatIsCancer?Availableonline: https://pancreatica.org/what-is-cancer/?gclid=Cj0KCQjwqc6aBhC4ARIsAN0NmMeDNTHgjo148-A5YsOvDAAuo0esSxivswW_WYxp1QpFAE3RU_BIhcaAuoyEALw_wcB

[9] Arabia, M.O.H.S. Cancer Facts and Guidelines. Available online: https://www.moh.gov.sa/en/HealthAwareness/EducationalContent/Diseases/Cancer/Pages/Cancer-2014-02-03-001.aspx

[10] GOWRISANTHYAM,DURGAKVand AKILAS"ColorectalCancer(CRC)Detection" QuestJournalsJournalofSoftwareEngineering and SimulationVolume10 ~Issue5 (2024)pp: 07-11ISSN(Online):2321-3795ISSN (Print):2321-3809www.questjournals.orgCRC2

[11] Md. Sanower Hossain, Hidayah Karuniawati,AmmarAbdulrahmanJairoun,Zannat Urbi,Der Jiun Ooi,Akbar John,Ya Chee Lim,K. M. Kaderi Kibria,A.K. M. Mohiuddin, Long Chiau Ming,Khang WenGoh and MuhammadAbdulHadi " Colorectal Cancer: A ReviewofCarcinogenesis,GlobalEpidemiology, CurrentChallenges,RiskFactors,Preventiveand Treatment Strategies"Cancers 2022, 14, 1732.https://doi.org/10.3390/cancers14071732,https://www.mdpi.com/journal/cancerHistopathological Classification of Colorectal Polyps using Deep Learning.

[12] MayPhuPaing,One-SunCho andJae-WanCho "Histopathological Classification of Colorectal PolypsusingDeep Learning"2023International ConferenceonInformationNetworking(ICOIN) |978-1-6654-6268-6/23/$31.00©2023IEEE| DOI: 10.1109/ICOIN56518.2023.10048925

[13] A. S. Sakr, N. F. Soliman, M. S. Al-Gaashani, P. Pławiak, A. A. Ateya, and M. Hammad, "An efficient deep learning approach for colon cancer detection," Applied Sciences, vol. 12, no. 17, p. 8450, 2022.

[14] B. Korbar et al., "Deep Learning for Classification of Colorectal Polyps on Whole-slide Images," J. Pathol. Inform., vol. 8, no. 1, p. 30, Jan. 2017, doi: 10.4103/jpi.jpi_34_17.

[15] Ahmed Sami Alhanaf, Saif Al-jumaili, Gokhan Bilgin, Adil Deniz Duru, Salam Alyassri and Hasan Huseyin Balik, "Deep Transfer Learning Methods for Classification Colorectal Cancer Based on Histology Images"2022 International Symposium on Multidisciplinary Studies and InnovativeTechnologies(ISMSIT)|978-1-6654-7013-1/22/$31.00©2022IEEE|DOI: 10.1109/ISMSIT56059.2022.9932746.

[16] Christian Szegedy1, Wei Liu, Yangqing Jia1, Pierre Sermanet1, Scott Reed,Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich "Going Deeper with Convolutions"

[17] Prashant Narayan Sholapur and Dr. IndirammaM "Explainable AI and Deep LearningtechniquesforColonCancerDetection"20 224th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)Inception.

[18] Vishwanath C. Burkapalli and Priyadarshini C. Patil,"TRANSFER LEARNING: INCEPTION-V3 BASED CUSTOM CLASSIFICATION APPROACH FOR FOOD IMAGES " ICTACT JOURNAL ON IMAGE AND VIDEO PROCESSING,AUGUST2020,VOLUME:11, ISSUE:01.

[19] Kaiming He,Xiangyu Zhang,Shaoqing Ren and Jian Sun, "Deep Residual Learning for Image Recognition".

[20] C. M. Bishop. Neural networks for pattern recognition. Oxford university press, 1995.