IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 4, April 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of **I**nnovative **R**esearch in **C**omputer and **C**ommunication **E**ngineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004140|

Lung Cancer Detection by CNN

Akshay Kalwaghe, Nutan Varpe, Pratibha Carpe

Department of Computer, Sahyadri Valley College of Engineering & Technology, Pune, India

ABSTRACT: Lung cancer is the cause of every sixth death around the world making it the second leading cause of death. Approximately 42 million people across the world suffer from cancer and this figure is continuously increasing. In India, approximately two and half million people are suffering from different types of cancer. If most of the cancers are detected in an early stage, then with the right remedy they can be cured. This paper offers details on a method in which CNN algorithm is used for the detection of illnesses like lung cancer. This paper also details the different machine learning techniques used to classify cancer into malignant and normal category. We get good accuracy on CNN which is 96.23% on 100 epochs.

KEYWORDS: Convolutional Neural Network, Deep Learning, Image processing

I. INTRODUCTION

Despite numerous developments in the field of diagnosis of illnesses like cancer, still, the tumor is one of the risky and dangerous illness. Lung Cancer is the second most popular cause for death not only in India but across the world. Diagnosis of a tumor is a totally critical and important task. The detection and remedy of cancerous tumors are one of the most important research and study areas. If the cancer is identified at an early stage and if right remedy is given quickly after the detection of the disorder, the rate of survival for the patients can be improved. There are numerous strategies or techniques used to seize various types of cancers, like PET scan, CT scan, Mammograms, MRI, 3D Ultrasound, Single Photon Emission Computed Tomography (SPECT), etc. Mammograms are used for breast cancer detection analysis. CT scan, MRI and several other techniques are used to identify brain tumors, lung cancer, etc.

The imaging method taken into consideration is mammogram and the type of classification strategies used are Feed forward back propagation, Extreme Learning Machine (ELM) ANN, backpropagation ANN, Particle Swarm Optimized Wavelet Neural Network, and CNN based on deep learning. For brain tumors, the imaging technique used is MRI and CT scan and the classification techniques considered are Level Set, K means Algorithm, SVM, Fuzzy C-means, Ad boost, Naïve Bayes classifier, and ANN classifier. For lung cancer, the medical imaging technique used is PET/CT. Also, classification techniques considered are FCM classifier, Feed Forward ANN, ANN, SVM binary classifier, and Entropy degradation method. Medical imaging techniques such as MRI and classification methods like ANN, SVM, and Multilayer perceptron neural network are considered for spine tumor detection. The two kinds of cancers are harmful and harmless growths. Standard MRI successions are for the most part used to separate various sorts of cerebrum cancers dependent on visual characteristics and different surface investigations of the delicate tissue. More than 120 classes of cerebrum cancers are known to be grouped in four levels as per the level of harm by the World Health Organization (WHO). A wide range of cerebrum cancers brings out certain indications dependent on the impacted district of the mind. The significant manifestations might incorporate migraines, seizures, vision issues, spewing, mental changes, memory slips, balance loss and so forth. Causes of cerebrum cancers are hereditary qualities, ionizing radiation cell phones, very low recurrence attractive fields, synthetic compounds, and head injury. Also, injury-resistant elements like infections and hypersensitivities may cause mild to severe cancer. The dangerous growths, otherwise called destructive cancers, are of two sorts - essential growths, which start from the cerebrum, and optional growths, which begin someplace and spread to the mind. The danger factors for mind growth are openness to vinyl chloride, neurofibromatosis, ionizing radiations, etc.

II. LITERATURE SURVEY

Wadood Abdul [1] used the architecture of CNN, a deep learning solution, in classifying the lung nodules as benign or malignant. LIDC-IDRI database was tested and the best results were obtained with 97.2% accuracy, 95.6% sensitivity, and 96.1% specificity, which outperforms the results obtained with other learning techniques. So, the ALCDC system performs better than the existing state-of-the-art systems.

Chao Ma, Gongning Luo, and Kuanquan Wang Waghmode et al. [2] stated that in this work, we introduce a new methodology that combines random forests and an active contour model for the automated segmentation of the glioma a type of tumor that occurs in the brain and spinal cord from multimodal volumetric MR images. Specifically, we employ a feature representation learning strategy to effectively explore both local and contextual information from multi-modal images for tissue segmentation by using modality-specific random forests as the feature learning kernels.

Onur Ozdemir et al. [3] proposed that the entirely 3D convolutional neural networks achieve state-of-the-art performance for both lung nodule detection and malignancy classification tasks on the publicly available LUNA16 and Kaggle Data Science Bowl challenges. It is important to have the coupling between detection and diagnosis components as nodule detection systems are typically designed and optimized on their own.

Anum Masood, Bin Sheng, Po Yang, and Ping Li, [4] proposed experimented enhanced multidimensional Region-based Fully Convolutional Network (mRFCN) based automated decision support system for lung nodule detection and classification. The mRFCN is used as an image classifier tool for feature extraction along with the novel multi-Layer fusion Region Proposal Network (mLRPN) with position-sensitive score maps (PSSM) being explored. They applied a median intensity projection to leverage three-dimensional information from CT scans and introduced a deconvolutional layer to adopt the proposed mLRPN in the architecture to automatically select potential regions of interest.

Khan Muhammad, Salman Khan [5] stated that an in-depth review of the surveys published so far and recent deep learning-based methods for BTC. Our survey covers the main steps of deep learning-based BTC methods, including pre-processing, features extraction, and classification, along with their achievements and limitations.

Chun-Mei Feng, Yong Xu [6] et al. Stated that discriminative information and sparsity in the PCA model. Specifically, in contrast to the traditional sparse PCA, which imposes sparsity on the loadings, here, sparse components are obtained to represent the data.

David N. Louis et al. [7] stated that notable changes include the addition of brain invasion as a criterion for atypical meningioma and the introduction of a soft tissue-type grading system for the new combined entity of solitary fibrous tumor human giopericytoma-a departure from how other CNS tumors are graded. Overall, this will facilitate clinical, experimental, and epidemiological studies that lead to improvements in the lives of patients with brain tumors.

Pär Salander et al. [8] proposed that most spouses witnessed months of global dysfunction preceding the symptom leading to physician consultation. The patient factors ‘less alien symptoms’, ‘personality change’ and ‘avoidance’; the spouse factors ‘spouse's passivity and ‘spouse's successive adaptation’; and the physician factors ‘reasonable alternative diagnosis’, ‘physician's inflexibility and ‘physician's personal values’ were identified as obstacles on the pathway to appropriate medical care.

Et al. [9] stated that the term brain tumor refers to a mixed group of neoplasms originating from intracranial tissues and the meninges with degrees of malignancy ranging from benign to aggressive. Each type of tumor has its biology, treatment, and prognosis and each is likely to be caused by different risk factors. Even benign tumors can be lethal due to their site in the brain, their ability to infiltrate locally, and their propensity to transform into malignancy. This creates problems in describing the epidemiology of these conditions and makes the classification of brain tumors a difficult science.

Jan J Heimans et al. [10] Proposed that a large number of Quality-of-Life instruments have been developed. The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) and the MOS Short-Form Health Survey are two frequently used general HRQL instruments. A specific brain tumor scale is the Brain Cancer Module, which is designed to be used in combination with general questionnaires. HRQL measurement and neuropsychological examination were used to investigate the impact of radiotherapy and surgery in low-grade glioma patients and the influence of tumor volume, tumor localization, performance status and age in both low-grade and high-grade glioma patients.

Malavika Suresh [11] stated that a noncognitive computer user interface has the endowment to perceive gestures and execute commands based on that. The design is implemented on a Linux system but can be implemented by installing modules for python on a windows system also. OpenCV and KERAS are the platforms used for identification. Gesture displayed on the screen is recognized by the vision-based algorithms. Using background removal technique, an assortment of skin color masks was trained by Lenet architecture in KERAS for the recognition.

M. Gurbina, M. Lascu, D. Lascu [12] stated that differentiate between a normal brain and a tumor brain (benign or malign). The study of some types of brain tumors such as metastatic bronchogenic carcinoma tumors, glioblastoma and sarcoma are performed using brain magnetic resonance imaging (MRI). The detection and classification of MRI brain tumors are implemented using different wavelet transforms and support vector machines. Accurate and automated classification of MRI brain images is extremely important for Medi-Cal analysis and interpretation.

S. Somasundaram, R. Gobinath [13] stated that focus on six features that are entropy, mean, correlation, contrast, energy and homogeneity. The performance metrics accuracy, sensitivity, and specificity are calculated to show that the proposed method is better compared to existing methods. The proposed technique uses MATLAB to detect the location and the size of a tumor in the brain through an MRI image.

Dhanasekaran Raghavan [14] proposed that the target with the aid of the following major steps, which include: Pre-processing of the brain images segmentation of pathological tissues Fluid (CSF)), extraction of the relevant features from each segmented tissue and classification of the tumor images with NN. As well, the experimental results and analysis are evaluated using Quality Rate (QR) with normal and abnormal Magnetic Resonance Imaging (MRI) images.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004140|

G. Hemanth; M. Janardhan [15] stated highly efficient and precise methods for brain tumor detection, classification and segmentation. To achieve this precise automatic or semi-automatic methods are needed. The research proposes an automatic segmentation method that relies upon CNN (Convolution Neural Networks), determining small 3 × 3 kernels. By incorporating this single technique, segmentation and classification are accomplished. CNN (an ML technique) from NN (Neural Networks) wherein it has layer-based for results classification. Various levels involved in the proposed mechanisms are data collection, pre-processing, average filtering, segmentation, feature extraction, CNN via classification and identification. With the use of Data Mining (DM) techniques, significant relations and patterns from the data can be extracted. The techniques of Machine Learning (ML) and DM are being effectively employed for brain tumor detection and prevention at an early stage of cancer.

S.K. Lakshmanaprabu [16] stated that Optimal Feature Level Fusion (OFLF) is considered to fuse low and high-level features of brain images; from this analysis, the images are classified as Benign or Malignant. From this implementation of medical images, the experiment results are evaluating performance metrics that are compared to existing classifiers. From the proposed MRI image classification process the accuracy was 96.23%, sensitivity was 92.3% whereas specificity was 94.52%; compared to the existing classifier. This proposed methodology is implemented in the working platform of MATLAB.

III. MATERIALS AND METHODS

A. Proposed Methodology

In a proposed system, we are proposing an experiment on lung cancer disease with a limited set of supervised data. We are proposing a combination of a Convolutional neural network-based multimodal disease risk prediction model with higher accuracy. We are going to solve the accuracy issue in the diagnosis of lung cancer with accurate stage predictions.

B. Dataset

We have collected the dataset from kaggle platform. We have split the dataset into two categories training and testing. For training 300 images and testing 60 mages we are used. Below is the following link

C. Pre-processing

In pre-processing we are convert the every image into 224*224.

D. Data augmentation

In data augmentation we are simply increase the dataset of training directory. We generate every image different format such as rotation, zoom and change the brightness of image.

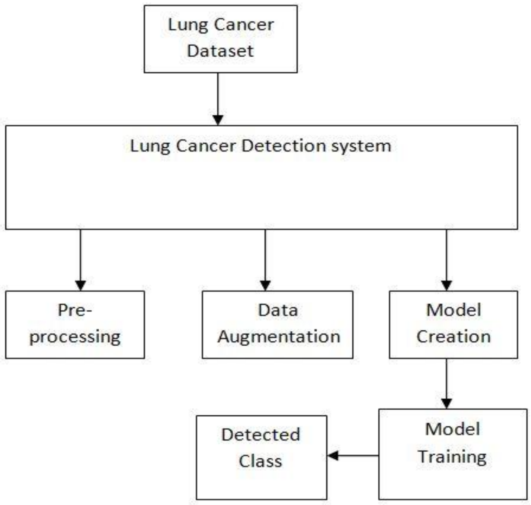

Fig. 1 Architecture Diagram

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004140|

E. Algorithms

1) Convolutional Neural Networks (CNN)

Convolutional Neural Networks (which are additionally called CNN/ConvNets) is a kind of Artificial Neural Networks that are known to be tremendously strong in the field of distinguishing proof just as picture order.

Four main operations in the Convolutional Neural Networks are shown in figure2.

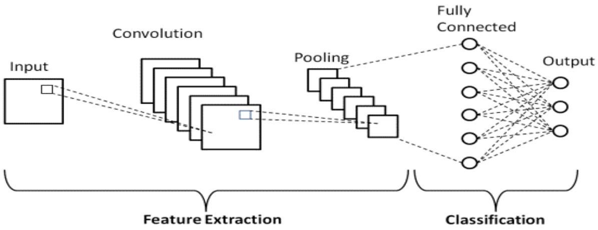

Fig. 2 Architecture of CNN

i) Convolution

This layer is the first layer that is used to extract the various features from the input images. Here, the mathematical operation of convolution is performed between the input image and a filter of a particular size MxM. By sliding the filter over the input image, the dot product is taken between the filter and the parts of the input image with respect to the size of the filter (MxM).

The output is termed the Feature map which gives us information about the image such as the corners and edges. Later, this feature map is fed to other layers to learn several other input image features.

ii) ReLU

ReLU (rectified linear unit) follows up on a basic level. It is an activity that is applied per pixel and overrides every one of the non-positive upsides of every pixel in the component map by nothing as shown in figure3.

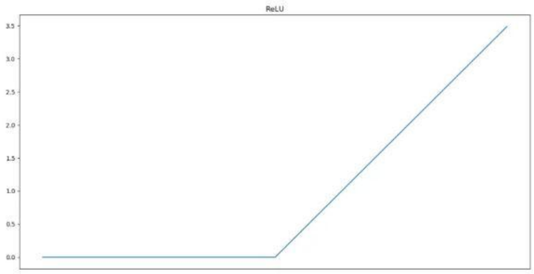

Fig. 3 Relu Activation

It is represented as:

$$f(x) = \begin{cases} 0, & if x<0 \\ x, & otherwise \end{cases} \quad \ldots\ldots\ldots\ldots (1)$$

iii). Pooling or sub-sampling

Spatial Pooling which is likewise called sub-sampling or downsampling helps in lessening the elements of each element map yet even at the same time, holds the most important data of the guide. Subsequent to pooling is done, in the long run, our 3D element map is changed over to a one-dimensional component vector.

iv) Fully connected

Neurons in layers are fully connected to all activations in the previous layer, as is the standard for feedforward neural networks. Fully Connected layers are always placed at the end of the network.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004140|

IV. RESULT AND DISCUSSION

Learning model is partially implemented. It is trained for 2 classes. For training and testing purpose total 150 images are used per class. Out of which 300 images are used for training dataset and 60 images are used for testing dataset. So 90% and 10% distribution is used for training and testing dataset respectively. Images are resized to 224*224 matrix and used as input to CNN. In our system we have used 128 filters for extracting the features from image. Once model is trained then we extracted the features from flatten layer and we have passed these features to fully connected layer. In the figure 6 and 7 we have shown the training graphs on X-axis epochs and y axis accuracy and loss. The epochs is increases the accuracy also increases as shown in figure7 vice versa in figure6 the epochs increases the loss decreases.

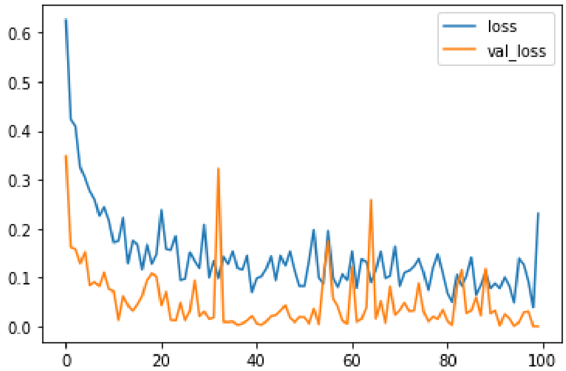

Fig.6 : Loss graph

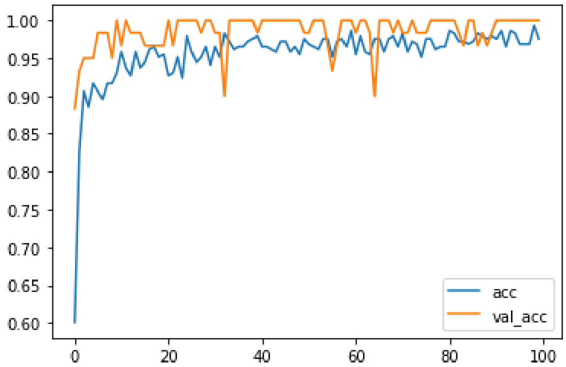

Fig.7: Accuracy graph

Sr. No.	Training Images	Testing Images	Algorithms	Accuracy (%)
1	300	60	CNN	96.23

Table1. Performance of Model

Once the CNN is trained with features of the training dataset, it performs the process of feature extraction in appropriate manner. The output of this trained CNN which performs the classification process. Batch size for training and testing are kept as 300 and 60 respectively. Training and testing accuracy of model for 100 epochs is shown in figure 4 and 5. It is observed that accuracy increases with number of epochs.

V. CONCLUSION

The proposed system proposes a hybrid approach for lung cancer detection systems over machine learning and CNN techniques. This system is a use of CNN algorithm which resolves the accuracy problem. The proposed system tries to improve accuracy and reduces the death rate. We get the 96.23% accuracy on 100 epochs. In future work we can implement on more diseases.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004140|

REFERENCES

[1] Wadood Abdul, "An Automatic Lung Cancer Detection and Classification (ALCDC) System Using Convolutional Neural Network", IEEE Transaction, 14 June 2021

[2] C. Ma, G. Luo, and K. Wang, "Concatenated and connected random forests with the multiscale patch is driven active contour model for automated brain tumor segmentation of MR images," IEEE Trans. Med. Imag., vol. 37, no. 8, pp. 1943–1954, Aug. 2018.

[3] C.-M. Feng, Y. Xu, J.-X. Liu, Y.-L. Gao, and C.-H. Zheng, "Supervised discriminative sparse PCA for com-characteristic gene selection and tumor classification on multiview biological data," IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 10, pp. 2926–2937, Oct. 2019.

[4] Onur Ozdemir, Rebecca L. Russell "A 3D Probabilistic Deep Learning System for Detection and Diagnosis of Lung Cancer Using Low-Dose CT Scans" IEEE TRANSACTIONS ON MEDICAL IMAGING last revised 21 Jan 2020.

[5] Anum Masood, Bin Sheng, Po Yang, Ping Li "Automated Decision Support System for Lung Cancer Detection and Classification via Enhanced RFCN with Multilayer Fusion RPN" IEEE Transactions on Industrial Informatics (Volume: 16, Issue: 12, Dec. 2020).

[6] Khan Muhammad; Salman Khan; Javier Del Ser; Victor Hugo C. de Albuquerque. "Deep Learning for Multigrade Brain Tumor Classification in Smart Healthcare Systems: A Prospective Survey" IEEE Transactions on Neural Networks and Learning Systems (Volume: 32, Issue: 2, Feb. 2021).

[7] David N. Louis, Arie Perry, et al. , "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary" , Acta Neuropathol , Springer May 2016

[8] Pär Salander, A Tommy Bergenheim, Katarina Hamberg, Roger Henriksson, "Pathways from symptoms to medical care: a descriptive study of symptom development and obstacles to early diagnosis in brain tumor patients", Family Practice, Volume 16, Issue 2, April 1999, Pages 143–148,

[9] McKinney PA ,"Brain tumors: incidence, survival, and etiology", Journal of Neurology, Neurosurgery & Psychiatry 2004;75:ii12-ii17.

[10] Heimans, J., Taphoorn, M. Impact of brain tumor treatment on quality of life. J Neurol 249, 955–960 (2002)

[11] Malavika Suresh, et al. "Real-Time Hand Gesture Recognition Using Deep Learning", International Journal of Innovations and Implementations in Engineering(ISSN 2454- 3489), 2019, vol 1

[12] M. Gurbină, M. Lascu and D. Lascu, "Tumor Detection and Classification of MRI Brain Image using Different Wavelet Transforms and Support Vector Machines", 42nd International Conference on Telecommunications and Signal Processing (TSP), Budapest, Hungary, 2019

[13] Somasundaram S and Gobinath R, "Early Brain Tumour Prediction using an Enhancement Feature Extraction Technique and Deep Neural Networks", International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278- 3075, Volume-8, Issue-10S, August 2019

[14] Damodharan S and Raghavan D, "Combining Tissue Segmentation and Neural Network for Brain Tumor Detection", The International Arab Journal of Information Technology, Vol. 12, No.1, January 2015

[15] G. Hemanth, M. Janardhan and L. Sujihelen, "Design and Implementing Brain Tumor Detection Using Machine Learning Approach", 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019

[16] A. R. Mathew and P. B. Anto, "Tumor detection and classification of MRI brain image using wavelet transform and SVM", International Conference on Signal Processing and Communication (ICSPC), Coimbatore, 2017.

[17] Nutnicha Teeyasuksaet "Image Classification of Rice Leaf Diseases Using random Forest", 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details