Health CPS Disease Prediction over Scanned Image using Machine Learning

Shubhangi Sarjerao Lomte[1] Snehal K. Lohar[2] Prajakta B. Mhaske[3] Sneha S. Sawant[4]
[1,2,3,4]D.Y.Patil College of Engineering and Technology, Kolhapur, India

Abstract— The field of medical imaging is gaining importance with an increase in the demand for automated, reliable, fast and efficient diagnosis which can provide insight to the image better than human eyes. Defects in brain is the second leading cause for cancer-related deaths in men in age 20 to 39 and fifth leading cause cancer among women in same age group. Defects in brain like tumors are painful and may result in various diseases if not cured properly. A prime reason behind an increase in the number of cancer patients worldwide is the ignorance towards treatment of a defected area in its early stages. The determination of defect extent is a major challenging task in brain defect planning and quantitative evaluation. Magnetic Resonance Imaging (MRI) is one of the non-invasive technique has emanated as a front- line diagnostic tool for brain defect without ionizing radiation. The automatic brain defect classification is very challenging task in large spatial and structural variability of surrounding region of defected area. In this work, automatic brain defect detection is proposed by using Convolutional Neural Networks (CNN) classification. The deeper architecture design is performed by using small kernels. Experimental results show that the CNN archives rate of 88-90% accuracy with low complexity and compared with the all other state of arts methods.

Keywords: Convolutional Neural Networks (CNN), Health CPS Disease Prediction, Machine Learning

I. INTRODUCTION

With big data growth in biomedical and healthcare communities, accurate analysis of medical data benefits early disease detection, patient care, and community services. However, the analysis accuracy is reduced when the quality of medical data is incomplete. Moreover, different regions exhibit unique characteristics of certain regional diseases, which may weaken the prediction of disease outbreaks. In this paper, we streamline machine learning algorithms for effective prediction of chronic disease outbreak in disease-frequent communities.

With Convolution neural network (CNN)-based multimodal disease risk prediction algorithm using structured and unstructured data from hospital we can analyse the disease. With the development of data analytics technology, more attention has been paid to disease prediction from the perspective of data analysis, various researches have been conducted by selecting the characteristics automatically from a large number of data to improve the accuracy of risk classification, rather than the previously selected characteristics. However, those existing work mostly considered structured data. For unstructured data, for example, using convolution neural network (CNN) to extract text characteristics automatically has already attracted wide attention and also achieved very good results. Furthermore, there is a large difference between diseases in different regions, primarily because of the diverse climate and living habits in the region.

Brain tumor segmentation is an important task in medical image processing. Early diagnosis of brain tumors plays an important role in improving treatment possibilities and increases the survival rate of the patients. Manual segmentation of the brain tumors for cancer diagnosis, from large amount of MRI images generated in clinical routine, is a difficult and time consuming task. There is a need for automatic brain tumor image segmentation.

Automatic segmentation using deep learning methods proved popular since these methods achieve the state-of-the-art results and can address this problem better than other methods.

II. NEED OF WORK

There is a need to design a system which will reduce the time needed to detect the disease from MRI image as we will analyse the disease or defect which will be easily recognized by doctor. This saves the time required for consultancy by analyser. With this system we can detect CPS related disease, so doctor will be able to notify patient regarding further treatment.

III. PURPOSE

The purpose is to detect the defect in brain from MRI images. The main reason for detection of defect in brain is to provide aid to clinical diagnosis. The aim is to provide an algorithm that guarantees to detect the defect in brain by combining several procedures to provide a foolproof method of brain defect detection in brain MRI images. The methods utilized are filtering, contrast adjustment, negation of an image, image subtraction, erosion, dilation, threshold, and outlining of the brain defect.

IV. MATERIALS AND METHODS

The research uses convolutional neural network for MRI brain defect segmentation using tensor flow. Normally, the segmentation is performed using various tools like MATLAB, LABVIEW etc. The research uses tensorFlow based MRI brain defect segmentation in order to improve segmentation accuracy, speed and sensitivity. Segmentation can be performed on MRI brain images and results are compared in terms of dice co-efficient.

A. *Requirements*

1) *Software Requirements:*

- Operating System: Ubuntu, Windows 8 and above
- Programming Language: Python V3.6
- Library: TensorFlow, Pandas, NumPy, Matplotlib, keras
- IDE : Anaconda Prompt

2) *Hardware Requirements:*

- Processor: Intel Core i5(8th gen) and above
- RAM: 8GB and above
- Graphics Card: 2GB

B. Python Based Convolutional Neural Network:

Many researchers use MAT LAB to implement the segmentation process. Our research work utilizes the python programming to implement the segmentation of MRI brain defect. The features are listed below in order to choose python programming to implement the research work

1) Python code is more compact and readable than MATLAB
2) The python data structure is superior to MATLAB
3) It is an open source and also provides more graphic packages and data sets

Hence, the proposed work utilizes python programming instead of MAT LAB. There are some additional python packages used during the implementation process of our research work through python.

V. PROPOSED WORK

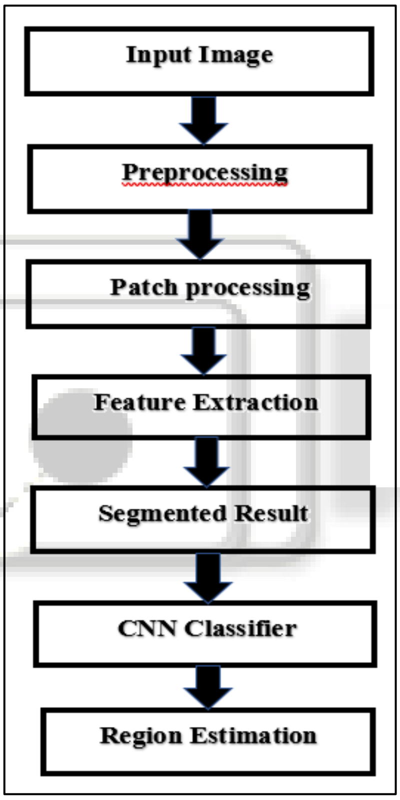

Fig. 1: Predictive Model of Proposed Work

A. MRI Brain Data Acquisition:

A brain MRI images dataset founded on Kaggle. The dataset contains 2 folders: abnormal and normal which contains 250 Brain MRI Images. The folder abnormal contains 150 Brain MRI Images that are defected and the folder normal contains 100 Brain MRI Images that are non-defected.

B. Preprocessing :

The main challenging task is removing artefacts produced by inhomogeneity in a magnetic field or small movements created by the patient during scanning. Many time bias is present in the scanning results, which affect the segmentation results, particularly in the computer n based models.

The bias correction removes the intensity gradient on each scanning images. Additionally, noise reduction is also performed by median filter in order to standardize the pixel intensities. Hence, noise reduction and bias correction helps to improve the data processing and provides the better segmentation, multiple radio frequency pulse sequences can be used to provide the different types of tissue.

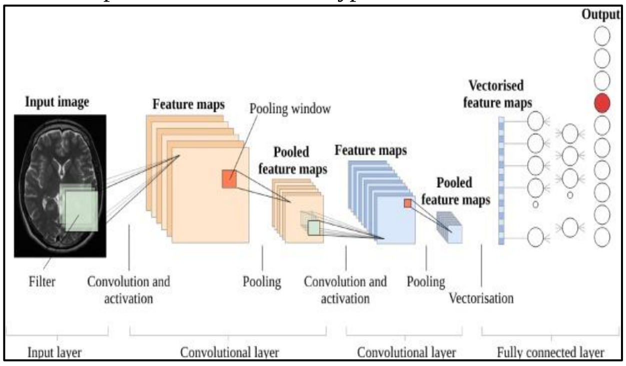

Fig. 2: The Architecture of Convolutional Neural Network

For every image, the following preprocessing steps were applied:

1) Crop the part of the image that contains only the brain
2) Resize the image to have a shape of (240, 240, 3) = (image_width, image_height, number of channels): because images in the dataset come in different sizes. So, all images should have the same shape to feed it as an input to the neural network.
3) Apply normalization: to scale pixel values to the range 0–1.

C. Data Augmentation:

Data augmentation is a strategy that enables practitioners to significantly increase the diversity of data available for training models, without actually collecting new data.
Before data augmentation, the dataset consisted of:
155 positive and 98 negative examples, resulting in 250 example images.
After data augmentation, now the dataset consists of:
1083 positive (53%) and 979 (47%) examples, resulting in 2062 example images.

D. Neural Network:

For image recognition, neural network is one of the powerful tools to perform segmentation. MRI is one of the most commonly used imaging techniques to capture MRI brain images. Automatic segmentation is a challenging task because of its large spatial and structural variability. Hence, the proposed system implements the automatic segmentation method based on CNN exploring small 3×3 kernals. The small size kernals help to design deeper architecture by using fewer number of weights in the network.

E. CNN Algorithm:

CNN algorithm performs the voxel-wise classification problem. The defected or lesion portion is separated from the background by calculating the probability of each image voxel belonging to the target is known. The CNN network has four sections input and convolution sections. The input layer processes the input image in order to produce the designed image patches. The convolution section process the designed image patches, in which multilayer convolutional filters operates and output feature maps. Further, the fully connected layer that groups all feature maps. The classification section estimates a prediction score to classify the every image voxel and provides a segmentation map.

F. Input Section:

The input section generates the image patches for the remaining of the network. The method performs the classification based on voxel-wise, where each voxel is classified based on the linear and nonlinear relationship between the focal voxel's intensity and its neighbours. The input 3D image size is large; hence, calculation of linear and nonlinear relationship between all voxels in the complete image is complex. Hence, the entire image is divided into smaller patches in order to find the relationship within a particular region instead of the entire image. It reduces the computational time and also memory space. It separates both local and global patches as input for the convolution section. For every extraction process, the central voxel is chosen randomly and extracted concentric local and global image patches. The neighbouring voxel around the central voxel provides the local information and global patch covers larger region and provides global information. In order to reduce the computation burden produced by the larger global patch we used down sampling of all global patches.

G. Convolution Section:

There are multiple layers in the convolution section, which help to sequentially identify the features using convolution operations. The captured features are low level features like edges and corner relationship between neighbouring voxels. The feature maps are the output of the convolutional layer. The convolutional layer calculates the output of the neurons that are connected to either local or global regions in the input. The convolution is the process of performing dot product between their inputs and their receptive field to which they are connected to in the input volume.

H. Data Split:

The data was split in the following way:

– 70% of the data for training.
– 15% of the data for validation (development).
– 15% of the data for testing.

VI. Experimental Results and Discussion

A. Procedure:

The experiment took place using following steps:

1) Step 1 :- Setting up Environment
2) Step 2 :- Data import And Pre-processing
3) Step 3 :- Training And Validation
4) Step 4 :- Data Augmentation
5) Step 5 :- Initializing Pretrain Model vgg16 architecture and weights as model
6) Step 6 :- Apply Augmentated data to the pretrain model
7) Step 7 :- Fetching vgg16 weights paths
8) Step 8 :- Running epochs to fit our model
9) Step 9 :- Save train model
10) Step 10 :- Loading keras and tensorflow
11) Step 11 :- Initializing our train model with parameters
12) Step 12 :- Passing Input to our train model
13) Step 13 :- Execution

B. Results:

Now, the best model (the one with the best validation accuracy) detects brain tumor with:
88.7% accuracy on the test set.
0.88 F1 score on the test set.
The results of the segmentation process are shown in following Fig.

Fig. 3: Input MRI Scanned Image

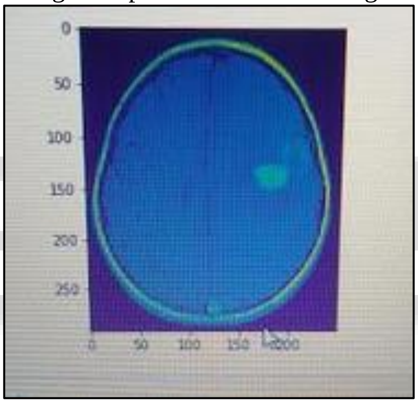

Fig. 4: Result of Brain Tumour Segmentation

VII. Conclusion

The CNN classification method is found to be more accurate with a percentage of 90, with increased sensitivity of 0.83 and higher specificity of 0.9342 in comparison with ANN method results.The CNN method found to be better than the ANN method in the any brain defect detection.

References

[1] MIN CHEN1 , (Senior Member, IEEE), YIXUE HAO1 , KAI HWANG2 ,(Life Fellow, IEEE), LU WANG1 , AND LIN WANG3,4 1" Disease Prediction by Machine Learning Over Big Data From Healthcare Communities" June 2017.
[2] P. B. Jensen, L. J. Jensen, and S. Brunak, ''Mining electronic health records: Towards better research applications and clinical care,'' Nature Rev. Genet., vol. 13, no. 6, pp. 395–405, 2012
[3] Buyue Qian Xiang wang,Nan Cao et al "A relative similarity based method for interactive patient risk prediction",DOI 10.1007/s10618-014-0379-5

