International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 03 Issue: 12 | Dec -2019 ISSN: 2582-3930

Lung Cancer Detection System Using Deep Learning

Rutuja Gholap,, Kajal Pisal, Sanchiti Phalle, Prerana Maske

Abstract—Deep Learning methods are becoming very popular in medical applications due to their high reliability and ease. Lung Cancer is reason for death of many people worldwide. Lung cancer detection is held by using various data analysis and classification methodologies. This review considers approaches to build automated diagnose system based on the lung cancer diagnose task. 3D convolution and recurrent neural networks are used to detect cancerous nodule.Finally, We are comparing the overall accuracy of the neural networks on lung cancer detection problem.

Keyword - deep learning, convolution neural network, lung cancer detection system

Index Terms—component, formatting, style, styling, insert

I. INTRODUCTION

Medical image process with automatic identification is one in every of the vital tasks which may improve the care eld and save ample patients life worldwide. It's the rationale why there are a great deal of efforts place into creation of the systems which may facilitate doctors to enhance performance and accuracy of the diagnosing. During this work, we have a tendency to take into account carcinoma detection drawback. Carcinoma may be a dangerous respiratory organ tumour featured by at liberty cell growth in tissues of the lung. Carcinoma patients have terribly low era within the advanced stages, wherever it's directly proportional with its growth at its detection time. There are multiple existing methodologies that are used for assessment of carcinoma like Computed Tomography (CT), Chest Radiography (x-ray), Magnetic Res-onance Imaging (MRI scan) and Humour Biology. Consistent with survey in 2018, it's ascertained that, The mean age for obtaining carcinoma : fifty four.6 years. The largest variety of carcinoma patients are over sixty five years more matured. Males predominate with a Male:Female magnitude relation of four.5:1 and this magnitude relation varies with age and smok-ing standing. The magnitude relation will increase increasingly up-to 51-60 years and so remains an equivalent. The smoker to non-smoker magnitude relation is high to 20:1 in varied studies. Lung Cancer New Cases Deaths Men forty eight,698 45,363 girls nineteen,097 18,112 each Sex sixty seven,795 63 475.

This paper involves sections. Section II and Section III gives brief introduction to Artificial Intelligence techniques namely Artificial Neural Network and Convolution Neural Network. Section IV involves detailed Literature Survey which presents the researches done by various authors in area of Lung Cancer Detection using above mentioned techniques.Section V TABLE I LUNG CANCER IN INDIA (GLOBOCAN 2018) [6]

Lung Cancer	New Cases	Deaths
Men	48,698	45,363
Women	19,097	18,112
Both Sex	67,795	63 475

presents problem definition for lung cancer detection system based on literature survey. In this paper Section VI defines AI techniques future scope.

II. ARTIFICIAL NEURAL NETWORK

Networks are nothing however the gathering of multiple nodes that releases biological neurons of

the human brain. Neurons are lay connected by links and that they act with one another. Nodules are accustomed take input and perform easy tasks on knowledge.

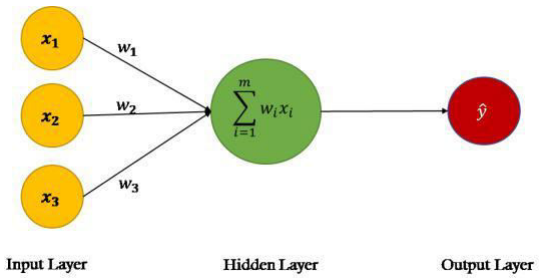

Fig. 1. Perceptron [8]

Perceptrons:

The perceptron is that the simplest neural network that con-sists of one somatic cell. Like Biological Neurons, perceptron could be a single articial somatic cell that consists of input nodes and single output node which is connected to every input node. Parts of Articial Neurons:

1) Input: Every input node is related to a numerical worth, which may be any imaginary number. Real numbers frame the complete spectrum of numbers: they'll be positive or negative, whole or decimal numbers.
2) Connections:
   Equally, every affiliation that departs from the input node includes a weight related to it, and this will even be any imaginary number.

3) Next, all the values of the input nodes and weights of the connections are brought together: they're used as inputs for a weighted sum: y=PD i=1 wixi , or, stated differently, y=f(w1x1+w2x2+...wDxD). This result are going to be the input for a transfer or activation perform.
4) Output:
   Output node is related to the perform (such because the sigmoid function) of the weighted add of the input nodes.
5) Bias:

Bias will be take into account because the weight related to an extra input node that's for good set to one. The bias worth is essential as a result of it permits you to shift the activation perform to the left or right, which may create a verify the success of your learning.

III. MAIN IMPORTANT DATA

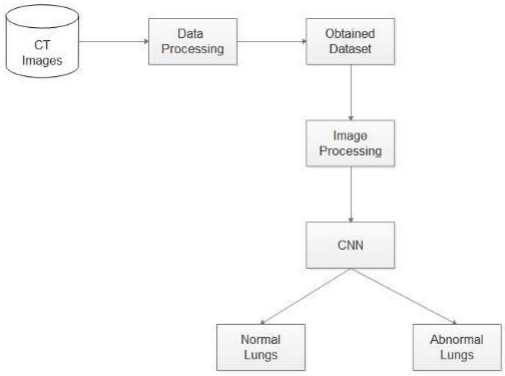

Fig. 2. System Architecture

We have planned our own model that depends on the reference paper we've got browse the general design of the carcinoma detection system is shown higher than. This carcinoma apprehension system consists of following steps:

1. Take a dataset of CT Scan Images.

Take a dataset of CT Scan pictures. Commencement is to require dataset that contains CT pictures of lungs. We have a tendency to are victimization Google Colaboratory setting for our project on that we've got uploaded our dataset. Our system contains DICOM dataset . There are virtually two hundred pictures in our dataset and it consist cancerous further as non-cancerous images which can facilitate North American country to simply diagnose the carcinoma.

2. Image Acquisition

In this step, we retrieve the images from our dataset to perform additional image pre-processing tasks. Image pre-processing is needed to coach, validate and take a look at our System model. 3. Image Pre-proccessing

Digital image process is nothing however the pc algorithms that perform image processing on digital

pictures. There are some image pre-processing techniques such as Noise Reduc-tion, Binarization.
Noise Reduction- Noise is nothing however the components within the image show variable pixel values rather than true pixel values that are obtained from actual image. Noise reduc-tion is that the method of decreasing the noise from the image. Perform a picture pre-processing to get rid of the noise the mathematician lter primarily based median ltering is provides higher results.
Binarization- Binarization is nothing however changing a component image to a binary image.
4. Body Region Extraction
During body region extraction stage, we have a tendency to are activity segmentation on the pictures to nd out region of interest. In segmentation, we have a tendency to divide the whole image into little small slices to envision that region has highest likelihood to urge malignant nodule.This task may be done by segmentation formula RCNN(Recurrent Convolution Neural Network). Segmentation facilitate North American country in detection of malignant nodule. Accuracy of nodule detection is extremely depends on operating of segmentation formula.
5. Feature Extraction
During this step, we have a tendency to extract the options from obtained region that we get from the segmentation. The options are perimeter, area, entropy, density, irregularity index, contrast, correlation etc.
6. Classification
Classication play major role in image process techniques. It's accustomed classify the options that are taken from the pictures. They're classied into numerous categories supported their totally different properties. Deep learning uses totally different classication techniques. During this system, we have a tendency to are victimization CNN classication technique as CNN itself distinguish categories with combination of algorithms.

IV. LITERATURE SURVEY

Bohdan Chapaliuk, Yuriy Zaychenko et.al.[1] (2018) has developed a model which based on two approaches. First is 3D convolution networks which segment 3D patient scan and classification of the most mallignacy segments. The second one is to use recurrent neural network to learn dependency between patient slices. There are two general approaches for distributed learning: model parallelism and data parallelism. There are different approaches to deal with neural networks parameters update.

Fausto Milletari[2], et.al.has proposed a volumetric fully convolution network for segmentation of 3D images. They have used MRI volumes of prostate to train the CNN and predict segmentation for the whole volume at once.They have introduced a special objective function that is optimised during training which is based on Dice Coefficient.They mainly deal with situations where there is a strong imbalance between foreground and background voxels. CNN is trained end-to-end on MRI volumes depicting prostate, and learns to predict segmentation for the whole volume at once.

Wentao Zhu[3], et.al.(2018)has proposed a fully automated CT cancer dignosis system based on deep learning called DeepLung. DeepLung consists of two parts, nodule detection and classi- fication. For nodule detection, they have designed 3D Faster R-CNN and U-net. Finally, gradient boosting machine with combined features are trained to classify nodules into malignant or non- malignant. The nodule classification subnetwork was validated on a public dataset from LIDC- IDRI, on which it achieved better performance than state-of-the-art approaches and surpassed the performance of experienced doctors based on image modality.

Fangzhou Liao[4], et.al.has evaluated the Malignancy of Pulmonary Nodules. The model consists of 2 modules. The first one is 3D region proposal network for nodule detection which gives all unsure nodules as output.The second one selects the top five nodules based on detection confidence and evaluate their cancer probabilities. The

overfitting caused by the shortage of the training data is alleviated by training the two modules alternately. The proposed model won the first place in the Data Science Bowl 2017 competition.

Rahul Dey[5], et.al. has developed a model which consists of two sequences GRU1 and GRU2 generated from the MNIST dataset and the IMDB dataset. The main driving signal of the gates appear to be the state as it contains essential information about other signals.It evaluates three variant models on MNIST and IMDB datasets of the Gated Recurrent Unit (GRU) in recurrent neural networks (RNNs) by retaining the structure and systematically reducing parameters in the update and reset gates.

Klaus Greff, Rupesh K. Srivastava[7] This paper reports the results of a large scale study on variants of the LSTM architecture. We conclude that the most commonly used LSTM architecture (vanilla LSTM) performs reasonably well on various datasets.A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell.

V. PROBLEM DEFINITION

All over the world among all types of cancer lung cancer is the main cause of death.The experienced doctors takes around 10 to 15 min to perform a detailed check for each patient because some nodules are small and hard to be found. Doctors can also evaluate the malevolence of nodule based on their structure, but the accuracy highly depends on doctors experience, and different doctors may give variable predictions. Computer Aided diagnosis is suitable for this task because computer vision models can quickly scan everywhere with equal quality and they are not affected by fatigue and emotions. Recent advancement of deep learning has enable computer vision models to help the doctor to diagnose various problems.

VI. CONCLUSION

It may be concluded from the above literature review that

There are two approaches which are used to detect lung cancer:

1. 3D Convolution neural network which is used for image processing i.e is segment 3D patient scan and classification of malignancy segments. 2.Use of recurrent neural network to learn dependency between patient slices. We can acheive human level accuracy on the different medical images by using these approaches. Result shows that neural network which are trained to detect lung cancer on whole lung 3D image gives worse accuracy in comparision to two stage approach, when two different neural network for segmentation and classification are trained. Recurrent neural network show competitive accuracy and performance.

VII. REFERENCES

1. Bohdan Chapaliuk, Yuriy Zaychenko ”Deep Learning Approach for Computer-Aided Detection Systems for Lung Cancer”,Volume 4, Issue 5, October 2018
2. Fausto Milletari ”V-Net: Fully Convolutional Neural Net-works for Volumetric Medical Image Segmentation”, June 2016
3. Wentao Zhu, ”DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification”, Jan 2018.
4. Fangzhou Liao, ”Evaluate the Malignancy of Pulmonary Nodules Using the 3D Deep Leaky Noisy-or Network”, Nov 2017
5. R. Dey, ”Gate-variants of Gated Recurrent Unit (GRU) neural networks,” 2017 IEEE 60th International Midwest Sym-posium on Circuits and Systems (MWSCAS), pp. 1597-1600, 2017.
6. http://cancerindia.org.in/lung-cancer/
7. K. Greff, R. K. Srivastava, ”LSTM: A Search Space Odyssey,” IEEE Transactions on Neural

Networks and Learn-ing Systems, vol. 28, no. 10, pp. 2222-2232, 2017.
8. https://images.app.goo.gl/LTsAqFKWrmVs5ruq7