International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 5 May 2021, pp: 141-146 www.ijaem.net ISSN: 2395-5252

Pulmonary Lobe Based Disease Prediction Using Deep Learning Techniques

V.Gokulakrishnan[1], T.Seenivasan[2], M.Velmurugan[3], S.Mishal Ahmed[4], R.Ragul[5]

[1]*Assistant Professor, Department of CSE, Dhanalakshmi Srinivasan Engineering College, Perambalur, Tamil Nadu, India.*

[2, 3, 4, 5] *UG Students, Department of CSE, Dhanalakshmi Srinivasan Engineering College, Perambalur, Tamil Nadu, India.*

Submitted: 01-05-2021 Revised: 09-05-2021 Accepted: 10-05-2021

ABSTRACT: Covid-19 is severe disease issue where a large number of people lose their lives every day. This disease affects not only a single country and even the whole world suffered this virus disease. In past decade, several kinds of viruses (like COPD, SARS, MERS and etc.) came into the picture. In the present time, the whole world is affected by Covid-19 disease and the most important thing is no single country scientists can prepare a vaccine for the same. COVID-19 is an infectious disease has spread all over the globe and is declared a pandemic. The detection of COVID-19 from chest X-ray and its differentiation from lung diseases with identical opacities is a puzzling task that relies on the availability of expert radiologists. Deep learning is an extremely powerful tool for learning complex, cognitive problems, to identify the diseases. In the present study, we have made use of a deep learning algorithm using the convolutional neural network (CNN) that can efficiently detect COVID-19 from chest X-ray images for swift diagnosis. Due to data scarcity related to COVID-19 chest X-ray images, instead of training the model from scratch, the present study made use of the Sparse matrix Construction are already available models in solving the analogous problems. Deep learning based classification models trained through the transfer learning approach can efficiently classify the chest X-ray images representing studied diseases. The analysis of this collected data is done with the help of CNN, a machine learning tool. This work mainly focuses on the use of CNN models for classifying chest X-ray images for coronavirus infected patients.

KEY WORDS - Convolutional neural network Algorithm (CNN), Histogram, Image processing, Segmentation.

I. INTRODUCTION

Deep learning is an artificial intelligence (AI) function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Deep learning is a subset of machine learning in artificial intelligence that has networks capable of learning unsupervised from data that is unstructured or unlabeled. Also known as deep neural learning or Artificial neural network. The artificial neural networks are built like the human brain, with neuron nodes connected together like a web.

The deep learning can involve with an image processing technique. In imaging science, image processing is processing of images using mathematical operations by using any form of signal processing for which the input is an image, a series of images, or a video, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Image processing usually refers to digital image processing but optical and analogue image processing also are possible.

II.RELEVANT WORK

An image processing is processing of images using mathematical operations by using any form of signal processing for which the input is an image, a series of images, or a video, the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it. Images are also processed as three-dimensional signals with the third-dimension being time or the z-axis. Image processing are used both analog and digital imaging. Traditional analog image editing is known as photo retouching, using tools such as editing tools. Raster images are stored in a computer in the form of a grid of picture elements, or pixels. These pixels contain the image's color and brightness information.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 5 May 2021, pp: 141-146 www.ijaem.net ISSN: 2395-5252

2.1SEGMENTATION

Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. The objective of the segmentation is to extract each character from the text present in the image. After performing Segmentation, the characters of the string will be separated and it will be used for further processing. The structure consists of a set of states plus transition probabilities between states.

Representation and Description

Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself. Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. Description deals with extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another.

Knowledge Base:

Knowledge may be as simple as detailing regions of an image where the information of interest is known to be located, thus limiting the search that has to be conducted in seeking that information.

Color based image segmentation:

Color segmentation may be more accurate because of more information at the pixel level comparing to gray scale images. The standard Red-Green-Blue (RGB) color representation has strongly interrelated color components and a number of other color systems have been designed. Choosing a proper color space is a very important issue for color image segmentation process. Generally L*A*B* and HSV are the two frequently chosen color spaces.

Texture based segmentation:

The texture analysis methods were used since the beginning of digital image processing. There are continuous suggestions for new algorithmic approaches to texture-based scene analyses.

Gabor Filter:

In image processing, a Gabor filter is a linear filter used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system and they have been found to be particularly appropriate for texture representation and discrimination.

Edge Detection:

Edges contain some of the most useful information in an image. We may use edges to measure the size of objects in an image; to isolate and to recognize particular objects from their background classify objects.

Shape based segmentation:

Segmentation results of the normalized cut to guide the shape model and thus avoid searching the shape space. First, we extract a shape space consisting of a mean shape and principal components by using PCA (Principal Component Analysis) from a training set. Second, we obtain the eigenvectors of Laplacian matrix derived from an affine matrix. Third, we project the segmentation generated from the eigenvectors onto the shape space to obtain a parametric shape model and then, based on the constraint of the shape model, we modify the similarities of pairwise pixels, which are elements of the affine matrix. Finally, we solve the modified Laplacian matrix again and update the shape model.

Parametric Shape Model:

There are several methods to represent the shape of an object, e.g. landmarks. They are followed by the set of features Binary images are sensitive to the boundary marked by hand in training samples

2.2 FEATURE EXTRACTION

In pattern recognition and in image processing, feature extraction is a special form of dimensional reduction. Transforming the input data into the set of features is called feature extraction. A feature is result of some calculations performed on the input data stream.

Color histograms:

The histograms are invariant to the rotation and translation and change only slowly under change of angle of view, change in scale and occlusion. Because histograms change slowly with view, a three dimensional object can be represented by a small number of histograms.

Edge Pixel Neighbourhood Information (EPNI):

In EPNI method the neighborhood edge pixels are found out that structure of those pixels will be used to make an extended feature vector.In computer vision and image retrieval process edge image matching is widely used for the comparison process.

Edge Histogram Descriptor (EHD):

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 5 May 2021, pp: 141-146 www.ijaem.net ISSN: 2395-5252

Edge Histogram Descriptor is the histogram generated using the edge pixels. The edge distribution is a good texture signature and also useful for image to image matching. This approach is not rotation invariant.

Angular radial partitioning (ARP):

In ARP method, the images in the stored database are converted to grayscale and edge detection is performed.

Eigenvector Approaches:

In eigenvector approach each image is represented by the small number of coefficients, which will be stored in the database and searched efficiently for the image retrieval and it is very successful for image retrieval.

SIFT-Scale Invariant Feature Transform:

Scale invariant feature transform (SIFT), as it transforms image data into scale-invariant coordinates relative to local features. An important aspect of this approach is that it generates large numbers of features that cover the image over the full range of scales and locations. A typical image of size 500×500 pixels will give rise to about 2000 stable features (although this number depends on image content).

Harris Corner:

Harris also showed its value for efficient motion tracking and the Harris corner detector has since been widely used for many other image matching tasks. While these feature detectors are usually called corner detectors.

Shape Descriptor:

Shape representation is mainly based on the shape features which are either based on the shape boundary information or boundary plus interior content.

Classification:

Classification includes a broad range of decision-theoretic approaches to the identification of images.

Neural Network Classification:

Image processing using Artificial neural networks (ANN) has been successfully used in various fields of activity such as geotechnics, civil engineering, etc. Image preprocessing, date reduction, segmentation and recognition are the processes used in managing images with ANN.

III.PROPOSED SYSTEM

Lung abnormality is one of the common diseases in humans of all age group and this disease may arise due to various reasons. Recently, the lung infection due to SARS-CoV-2 has affected a larger human community globally and due to its rapidity, the World health organization (WHO) declared it as pandemic disease. The COVID-19 disease has adverse effects on the respiratory system and the infection severity can be detected using a chosen imaging modality. The field of medical imaging introduced CAD (Computer aided diagnostic) systems which help medical specialist to identify and categories the problem. The lesions are produced with different body parts which cause the cancer. The next task is to separate the objects in lungs volume for non-part of lungs. These objects are unwanted lesions. The next step is to classify the potential nodules into nodules and non-nodule. The lung lobes contain holes, which are analysed using active contour method. The contour correction is performed to include the fissure nodule. The candidate nodules are extracted using different levels of probabilistic lobe segmentation. The candidate nodules are pruned using group fissure prior. The hybrid features are extracted from pruned candidate nodules. Based on these features, classify the lung lobes to predict the lung diseases. In the proposed work; COVID-19 is detected using deep learning from CT scan images decomposed to sparse matrix. A deep learning based detection model is proposed to improve the detection accuracy.

STEPS:

Image Acquisition

Computed Tomography (CT) is taken into account in concert of the simplest strategies to diagnose the pneumonic nodules. It uses x-rays to get structural and practical info concerning the physical body. However, the CT image quality is influenced lots by the radiation dose. The standard of image will increase with the many quantity of radiation dose, however within the same time, this will increase the amount of x-rays being absorbed by the lungs. In this architecture contains four phases such as pre-processing, segmentation, features extraction and classification.

Pre-Processing

Here, CT image can input to the system. The user has to select the required lung frame image for further processing. Then each image is resized to 256*256. Then implement median filter to remove noises from lung images. The median filter is a nonlinear digital filtering technique, often used to remove noise from an image or signal. Such noise

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 5 May 2021, pp: 141-146 www.ijaem.net ISSN: 2395-5252

reduction is a typical pre-processing step to improve the results of later processing. Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise (but see discussion below), also having applications in signal processing. The main idea of the median filter is to run through the signal entry by entry, replacing each entry with the median of neighbouring entries. Median filtering is a nonlinear method used to remove noise from images. It is particularly effective at removing 'salt and pepper' type noise. The median filter works by moving through the image pixel by pixel, replacing each value with the median value of neighboring pixels. The pattern of neighbours is called the "window", which slides, pixel by pixel over the entire image pixel, over the entire image.

Features Extraction

Feature learning comprises a set of algorithms to transform labeled or unlabeled data to a new space, where it can capture the parameters and patterns of variation by disentangling the hidden features. Features are learned through supervised and unsupervised learning scheme. Numerous unlabeled data is available in each domain, e.g. images, text data, speech, which contain several patterns of variation that can easily be collected for feature extraction, e.g. from pre-processed image. The task of feature extraction from unlabeled data is known as unsupervised feature learning. Data-adaptive representations are dependent on the statistics of data. Such representations are learned directly from the observed data by optimizing some measure that quantifies the desired properties of the representation. In linear sparse coding, the goal is to find a decomposition in which the hidden components are sparse, meaning that they have probability densities which are highly peaked at zero and have heavy tails.

Lung Classification

The classification is the final step of the system. After analysing the structure, each section individually evaluated for the probability of true positives. Lung diseases are classified using Convolutional neural network algorithm. CNNs represent feed-forward neural networks which encompass diverse combos of the convolutional layers, max pooling layers and completely related layers and Take advantage of spatially neighbourhood correlation by way of way of imposing a nearby connectivity pattern among neurons of adjacent layers. A CNN includes one or extra pairs of convolution and max pooling layers and ultimately ends with completely related neural networks. The hierarchical structure of CNNs is steadily proved to be the most efficient and successful manner to analyse visible representations. We can see that the curve of every class has its own visual shape which is different from other classes, although it is relatively difficult to distinguish some classes with human eye (e.g., gravel and self-blocking bricks).

Disease Prediction

Covid-19 is a rapidly spreading viral disease that infects not only humans, but animals are also infected because of this disease. The daily life of human beings, their health, and the economy of a country are affected due to this deadly viral disease. Covid-19 is a common spreading disease and till now, not a single country can prepare a vaccine for COVID-19. A clinical study of COVID-19 infected patients has shown that these types of patients are mostly infected from a lung infection after coming in contact with this disease. Chest x-ray (i.e., radiography) and chest CT are a more effective imaging technique for diagnosing lunge related problems. Still, a substantial chest x-ray is a lower cost process in comparison to chest CT. In this module, we can identify the COVID-19 and other diseases. And also provide prescription for affected diseases.

IV. SYSTEM DESIGN

This architecture has five parts such as pre-processing, features extraction, segmentation and classification. User can be input Lung CT image as input and pre-processing steps to convert the image into grey scale and filter the noise using median filter algorithm. Then perform features extraction steps to extract the color, shape and other features and constructed as sparse matrix. After that perform active contour method to segment the lung boundaries. Finally classify the features whether is affected or not.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 5 May 2021, pp: 141-146 www.ijaem.net ISSN: 2395-5252

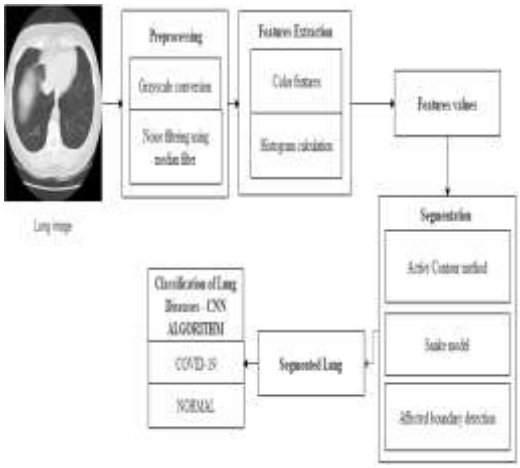

Fig 4.1 System Architecture

V.SYSTEM TESTING

Software testing is a method of assessing the functionality of a software program. There are many different types of software testing but the two main categories are dynamic testing and static testing. Dynamic testing is an assessment that is conducted while the program is executed; static testing, on the other hand, is an examination of the program's code and associated documentation. Dynamic and static methods are often used together.

Testing is a set activity that can be planned and conducted systematically. Testing begins at the module level and work towards the integration of entire computers based system. Nothing is complete without testing, as it is vital success of the system.

There are three ways to test a program

- For Correctness
- For Implementation Efficiency
- For Computational Complexity.

Tests for correctness are supposed to verify that a program does exactly what it was designed to do.

The data is entered in all forms separately and whenever an error occurred, it is corrected immediately. A quality team deputed by the management verified all the necessary documents and tested the Software while entering the data at all levels. The development process involves various types of testing. Each test type addresses a specific testing requirement. The most common types of testing involved in the development process are:

- Unit Test.
- Functional Test
- Integration Test

Unit Testing

The first test in the development process is the unit test. The source code is normally divided into modules, which in turn are divided into smaller units called units. These units have specific behaviour. The test done on these units of code is called unit test.

Functional Testing

Functional test can be defined as testing two or more modules together with the intent of finding defects, demonstrating that defects are not present, verifying that the module performs its intended functions as stated in the specification and establishing confidence that a program does what it is supposed to do.

Integration Testing

In integration testing modules are combined and tested as a group. Integration Testing follows unit testing and precedes system testing. Testing after the product is code complete.

VI.CONCLUSION

The confirmatory diagnosis of COVID-19 is mainly dependent on clinical symptoms, epidemiological history, nucleic acid detection, immune identification technology, etc. All the methods mentioned above have some limitations such as time required, costs, equipment dependence, shortage of testing kits, availability of trained healthcare workers, inter-operator variability's, especially in a pandemic like this, making them cumbersome diagnostic procedures. Timely diagnosis of the COVID-19 patients can enable help in the optimization of available resources, including trained human resources, for all the supportive measures required for confirmed patients. Automated AI-based intelligent chest X-ray classification has such untapped potential for this unmet need, as evident from recent researches. Rapid screening to diagnose such patients is also essential for controlling outbreaks. In conclusion, an AI system derived from heterogeneous multinational training data delivers acceptable performance metrics for the classification of chest CT for COVID-19 infection. While CT imaging may not necessarily be actively used in the diagnosis and screening for COVID-19, this deep learning-based AI approach may serve as a standardized and objective tool to assist the assessment of imaging findings of COVID-19 and may potentially be useful as a research tool, clinical trial response metric, or perhaps as a complementary test tool in very specific limited populations or for recurrent outbreaks settings.

REFERENCES

[1]. A. S. Fauci, H. C. Lane, and R. R. Redfield, "Covid-19 - navigating the uncharted." The New England journal of medicine, vol. 382, pp. 1268–1269, Mar. 2020.

[2]. N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, J. Xia, T. Yu, X. Zhang, and L. Zhang, "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in wuhan, china: a descriptive study." Lancet (London, England), vol. 395, pp. 507–513, Feb. 2020.

[3]. W.-J. Guan, Z.-Y. Ni, Y. Hu, "Clinical characteristics of coronavirus disease 2019 in china." The New England journal of medicine, vol. 382, pp. 1708–1720, Apr. 2020.

[4]. D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, and Z. Peng, "Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in wuhan, china." JAMA, Feb. 2020.

[5]. H. Shi, X. Han, N. Jiang, Y. Cao, O. Alwalid, J. Gu, Y. Fan, and C. Zheng, "Radiological findings from 81 patients with covid19 pneumonia in Wuhan, china: a descriptive study." The Lancet. Infectious diseases, vol. 20, pp. 425–434, Apr. 2020.

[6]. Xie, Xingzhi, et al. "Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing." Radiology 296.2 (2020): E41-E45.

[7]. Mei, Xueyan, et al. "Artificial intelligence–enabled rapid diagnosis of patients with COVID-19." Nature medicine 26.8 (2020): 1224-1228.

[8]. Shi, Feng, et al. "Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19." IEEE reviews in biomedical engineering (2020).

[9]. Irvin, Jeremy, et al. "Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. No. 01. 2019.

[10]. Johnson, Alistair EW, et al. "MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs." arXiv preprint arXiv: 1901.07042 (2019).