IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Enhanced Diagnostic Accuracy For Pneumonia Using Machine Learning On X-Ray Images

Mrs. B. Sathyabama
Assistant Professor
Department of Computer Applications
T. John college, Bangalore, Karnataka, India

Abstract: 'Pneumoscan' is a web application developed to diagnose pneumonia on chest X-ray and medical parameters using the concept of deep learning algorithms. Inthis model pneumonia and non-pneumonia condition can be detected using CAAD model which has an anomaly detection module and confidence prediction module, which shares a feature extractor. If the anomaly score is higher than the threshold or confidence score is small than the threshold, the entered input will be treated as apneumonia case and we will compare this with the medical parameter and give the severity level. The system is implemented using Django framework and Python.

Keywords: CAAD, Radiology, Image Acquisition, CNN Model.

INTRODUCTION

The main aim of this paper is the screening of viral pneumonia on chest X-ray using CAAD (Confidence-Aware Anomaly Detection) model. The project addresses the limitations of traditional pneumonia screening methods such as manual interpretation by radiologists or computer aided detection systems, which can be time-consuming and subject to human error. CAAD model, have the potential to overcome these limitations and improve the accuracy and efficiency of screening for pneumonia on chest X-ray. The system also has the feature to input multiple images, so that the doctor, radiologists, or patient can provide the system with multiple inputs. The projects proposed method for viral pneumonia screening on chest X-rays using CAAD model has significant implications for the diagnosis of covid patients and could contributeto the global effort to manage the pandemic.

OBJECTIVE

The main objective was to provide a space that help the patients, doctor, radiologists to examine X-ray and conclude that whether the person is having pneumonia or not. This system will help to reduce the errors caused by individual while making an observation, it will increase the efficiency.

The paper aims to train and test the CAAD model on a large dataset of chest X-rays from patients with viral pneumonia and evaluate its potential for improving the accuracy and efficiency of pneumonia screening. The system can also be used to detect covid-19 pneumonia as well as other type of viral pneumonia.

PROBLEM IDENTIFICATION

Presently, pneumonia detection relied on human expertise and interpretation of the images. Additionally, with the onset of the COVID-19 pandemic, there has been an urgent need for more accurate and efficient screening for viral pneumonia. Therefore, there is a need for a more advanced and reliable system for detecting viral pneumonia on chest X-rays. The goal of this project is to detect pneumonia from a chest X-ray by using CAAD (Confidence Aware Anomaly Detection) which solved the issues of human error or misinterpretation, as well as a lack of consistency and standardization in diagnosis.

Gantt chart

The Gantt chart is an excellent tool for quickly assessing the status of a project. It is a project control technique that can be used for several purposes. It is used for tracking and reporting progress, as well as for graphically displaying a schedule. Gantt chart is used to report progress they represent in easily understood picture & project status.

It includes:

- Scheduling
- Budgeting
- Resource planning

This allows us to see at a glance:

- What the various activities are.
- When each activity begins and ends.
- How long each activity is scheduled to last.
- Where activities overlap with other activities, and by how much.
- The start and end date of the whole project.

LITERATURE REVIEW

This section covers a survey of the most recently published disease diagnosis techniques. Several disease detection systems have been proposed which vary in terms of effectiveness and robustness.

(i) "Anomaly Detection in Medical Imaging -- A Mini Review (Published: 25 Aug 2021)" The authors describe the anomaly detection is the task of identifying unusual samples from the majority of the data. This paper tries to provide a structured and comprehensiveoverview of the research on anomaly detection. They grouped existing techniques into different categories based on the basic approach adopted by each technique. For each category they have identified key assumptions, which are used by the techniques to differentiate between normal and anomalous behavior.

(ii) "Radiologic diagnosis of patients with COVID-19 (Published: February 2021)"

The authors describe finding pneumonia disease by CNN concepts. A chest X-ray used to diagnose pneumonia should be evaluated by an experienced radiation oncologist. Therefore, the development of an automatic pneumonia detection system would be beneficial for prompt treatment of the disease. Due to the success of deep learning algorithms in analyzing medical images, Convolutional Neural Networks (CNNs) have gained much attention for disease classification. In addition, features learned by pre-trainedCNN models on large-scale datasets are much useful in image classification tasks. Chest X-rays, which are used to diagnose pneumonia, need expert radiotherapists to evaluate. The evaluation of chest X-Ray specifically in case of Pneumonia can be misleading because many other problems like congestive heart failure, lung scarring etc. can mimic a Pneumonia. This is the main reason for the misclassification of X-ray images in our data set. Thus, it is challenging to develop an algorithm for detecting thoracic diseases like Pneumonia. They evaluated the performance of different variants of pre-trained CNN models followed by different classifiers for classifying abnormal and normal chest X-Rays.

(iii) "A review on deep learning in medical image analysis (Published: 04 September 2021)" The authors describe A review on deep learning in medical image analysis. Computer processing and analysis of medical images includes image acquisition, image generation, image analysis, and image-based visualization. Medical imaging has evolved into computer vision, pattern recognition, image mining, and various strands of machine learning. Deep learning is a commonly used method to ensure accuracy. Deep learning uses neural networks consisting of many convolutional nodes of artificial neurons to learn patterns in data structures.

(iv) Pre-processing methods in chest X-ray image classification (Published: April 5, 2022)"The authors provide proof that machine learning can be used to support medics in chest X- ray classification and improving pre-processing leads to improvements in accuracy, precision.

This paper describes that the patient with Covid - 19 has abnormalities in chest - X rays. Chest - X rays are more available than CT images. Chest - x-ray analysis can be time-consuming and requires highly educated specialists to interpret. But, the use of machine learning (ML)-based methods can improve efficiency, support medics in the diagnosis of COVID-19, and speed up the time to diagnosis.

(v) "Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19 (Published Online: Mar 27 2020)" The authors describe deep learning of chest radiography for finding pulmonary tuberculosis by using convolution neural networks. Using deep learning on chest X-ray images, task categorization works: image-level prediction (classification and regression), segmentation, localization, image generation. The aim of the paper is to evaluate the efficacy of deep convolutional neural networks (DCNNs) for detecting tuberculosis (TB) on chest radiographs. Deep learning using convolutional neural networks can accurately classify tuberculosis on chest x-rays.

EXISTING SCHEME

In the existing scheme, we relied primarily on human expertise and interpretation. Medical professionals, such as radiologists or doctors would analyze the image to identify signs of pneumonia, including the presence of infiltrates, consolidation or nodules. However, the process could be time-consuming and potentially prone to errors, as different professionals may have varying levels of experience and training. Interpreting chest X-ray manually can be time consuming, especially in cases where the doctor needs to review multiple images over time to track changes, during pandemic situation it resulted as a problem. The manual interpretation is costly.

Disadvantages

- The interpretation of Xray is subjective, can vary depending on experience.
- The manual process is time consuming.
- Diagnosis is error-prone
- Costly.

PROPOSED SCHEME

In the proposed scheme an advanced Deep-learning based diagnostic system CAAD model is designed to accurately and efficiently detect viral pneumonia on chest X-ray. The CAAD model combines deep learning algorithm with anomaly detection technique to analyze the X-ay. The system is designed to provide healthcare professionals with a fast, reliable, and consistent approach to the diagnosis of pneumonia. This system is significant because it has the potential to improve the accuracy and efficiency of screening for viral pneumonia, which can be difficult to diagnose using traditional methods. The feature of uploading multiple image feature is also associated with

this system it will help to tackle the pandemic situation. The system has thepotential to reduce healthcare cost associated with the diagnosis of pneumonia.

Advantages

- Accurate and efficient diagnosis.
- Consistent and reliable diagnosis.
- Reduced Subjectivity.
- Secure and cost effective.

SYSTEM ARCHITECTURE

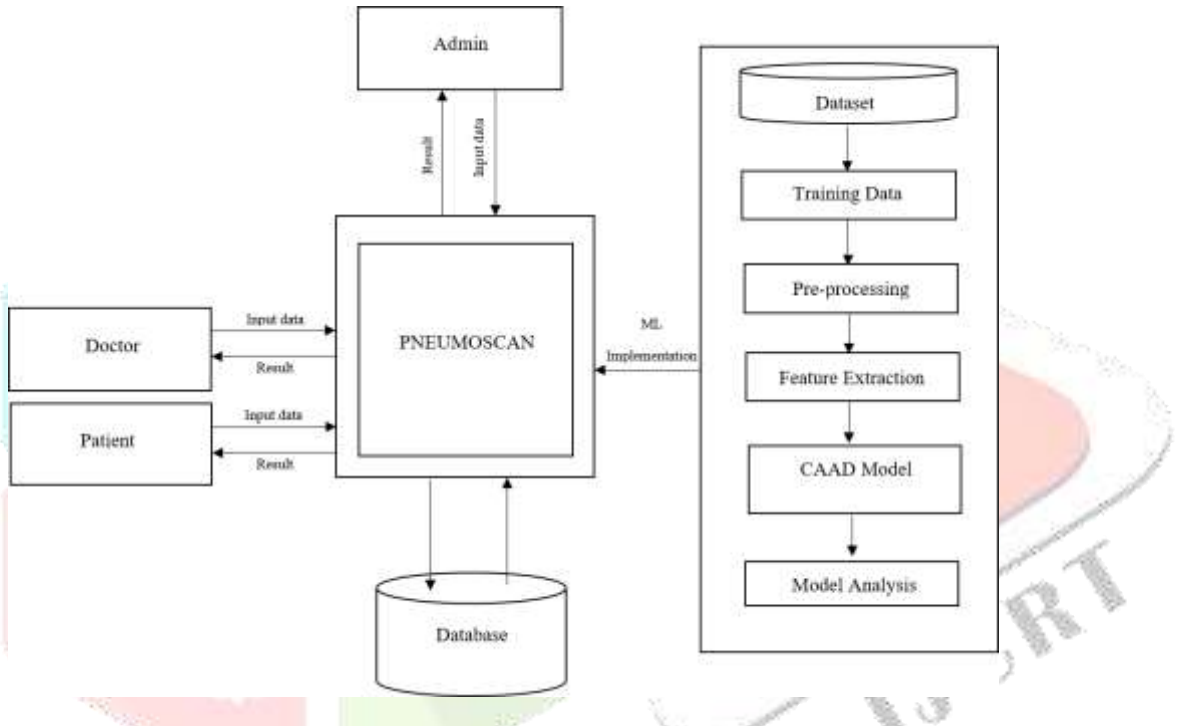

Fig: System Architecture of Pneumoscan

METHODOLOGIES USED:

Dataset

A Kaggle dataset contains chest x-ray images of Covid-19, Pneumonia and normalpatients. Dataset is organized into 2 folders (train, test) and both train and test contain 3 subfolders (COVID19, PNEUMONIA, NORMAL). Dataset contains total 6432 x-ray images and test data have 20% of total images.

Image Pre-processing

Image pre-processing is the process of implementing various method such as resizing the images, converting them to suitable format, and the augmentation of images for enabling the digital image data to be used within the machine learning model effectively.

Training

After completing data pre-processing, the inception v3 model will be created which consists of various convolutional and pooling layers alongside a 1x1, 3x3& 5x5 sized kernels. The model will be then trained on the basis of training and validation data with the help of several python libraries that where preloaded to perform this specific tasks.

Testing Accuracy

After the model is trained using the training dataset, we use the testing dataset to evaluate how well it works. A particular part of the overall dataset is used as the testing data set on the basis of which the accuracy is computed for the proposed model and an accuracy above 95% is obtained.

Uploading image by users

After the model is evaluated i.e., trained and tested using the dataset, the trained modelis then used by the user, where he/she uploads images using the mouse cursor by clicking the mouse.

Disease diagnosis

After the user upload images, the system will recognize the diseases using the particular models.

Result

The result of the diagnosis will be shown to the user.

FEASIBILITY STUDY

The feasibility study is major factor, which contributes to the analysis and development of the system. The decision of the system analyst whether to design a particular system or not depends on its feasibility study.

Operational Feasibility:

- The introduction to this system is not going to hamper any user of the system.
- The proposed system is very flexible and user friendly. The proposed system produces best results and gives high performance. It can be implemented easily. So, this project is operationally feasible.

Technical Feasibility:

- This feasibility deals with technicality of the system. Instead of performing stress analysis in a traditional way, this system is fully automated.
- No efficient workers are required to handle the system.

Economic Feasibility:

Economic Feasibility deals about the economic impact faced by the organization toimplement a new system.

- The cost to conduct a full system investigation is possible.
- There is no additional manpower requirement.
- There is no additional cost involved in maintaining the proposed system.

SYSTEM IMPLEMENTATION

CAAD MODEL

The model is based on a convolutional neural network architecture which is a type of neural network commonly used for image analysis. The method works by training a deep learningmodel to identify normal chest Xray and the using the information to detect anomalies that may indicate of pneumonia. In this model, it is trained on a large dataset of chest X-rays that are labeled to indicate whether the image is normal or abnormal. Once the model has learned how normal is, it can then be used to detect anomalies or deviations from the learned normal representation.

To detect anomalies in chest x-ray that may be indicative of pneumonia, themodel is fed with an input image and if the input image is similar to the learned normal representation the case is considered as normal. If the input image contains anomalies opacities or consolidations in the lungs that are indicative of pneumonia. To accomplish this, the anomaly detection module uses a threshold-based approach where it compares the distance between the normal X-ray and the input X-ray to a predefined threshold value. If the distance is above the threshold the input X-ray is classified as abnormal. The threshold value is determined during the training phase and is set based on the distribution of distance between the learned representation of normal X-ray and the training dataset. This distance is a measure of how different the input X- ray is from the what the model has learned.

1.Data preprocessing: Preprocessing the chest X-ray image by resizing them to a uniform size of 256*256 pixels and normalizing the pixel values.

2. Training the CNN model: The system train CNN model on a large dataset of normal and abnormal chest X-rays. The CNN model is trained to extract the relevant features from the input images and classify them as normal or abnormal.

3. Anomaly detection module: To detect anomalies in chest x-ray that may be indicative of pneumonia, the model is fed with an input image and if the input image is similar to the learned normal representation the case is considered as normal. If the input image contains anomalies opacities or consolidations in the lungs that are indicative of pneumonia. To accomplish this, the anomaly detection modules use a threshold-based approach where it compares the distancebetween the normal X-ray and the input X-ray to a predefined threshold value. If the distance is above the threshold the input X-ray is classified as abnormal. The threshold value is determined during the training phase and is set based on the distribution of distance between the learned representation of normal X-ray and the training dataset. This distance is a measure of how different the input X-ray is from the what the model has learned.

4. Confidence prediction module: The system uses a confidence prediction module to estimate the confidence score in its predictions for each input image. This confidence prediction is used to select highly confident anomaly samples for training the anomaly detection module.

CONCLUSION

'Pneumoscan' utilizes a convolutional neural network (CNN) for feature extraction and classification, followed by an anomaly detection module and a confidence prediction module for identifying highly confident anomaly samples. The experimental results showed that the proposed method achieved high accuracy and sensitivity on a large and diverse dataset of chest X-ray images, outperforming several manual methods for viral pneumonia screening. The method also demonstrated good generalization performance on a separate test dataset, indicating its potential for use in real-world clinical settings

FUTURE SCOPE

- The system can be integrated with Electronic Health Records to provide real-time screening and diagnosis of viral pneumonia. This integration can help doctors and medical practitioners to quickly identify cases of viral pneumonia and initiate prompt treatment.
- The system can be evaluated on larger and more diverse datasets to assess its performance on a wider range of populations.
- The system can be extended to incorporate other imaging modalities, such as CT scans, to improve the accuracy of the diagnosis and reduce the false positives. The system can also be extended to the detection of other diseases or conditions on chest X-rays, such as tuberculosis or lung cancer.
- The system can be further improved by integrating it with more advanced AI algorithms and

machine learning models in future, which can help enhance its accuracy and performance.

BIBLIOGRAPHY

[1] Y. Bai, L. Yao, T. Wei, F. Tian, D.-Y. Jin, L. Chen, and M. Wang, "Presumed asymptomatic carrier transmission of COVID-19," Jama, vol. 323, no. 14, pp. 1406–1407, 2020.

[2] H. Chen et al., "Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: A retrospective review of medical records,"Lancet, vol. 395, no. 10226, pp. 809–815, Mar. 2020.

[3] W. Li et al., "Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus," Nature, vol. 426, no. 6965, pp. 450–454, 2003.

[4] E. I. Azhar et al., "Evidence for Camel-to-Human transmission of MERS coronavirus," New England J. Med., vol. 370, no. 26, pp. 2499–2505, Jun. 2014.

[5] V. M. Corman et al., "Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR," Eurosurveillance, vol. 25, no. 3, Jan. 2020, Art. no. 2000045.

[6] L. Lan et al., "Positive RT-PCR test results in patients recovered from COVID-19," J. Amer. Med. Assoc., vol. 323, no. 15, pp. 1502–1503, 2020.