HBRP PUBLICATION

Journal of Advancement in Parallel Computing
Volume 6 Issue 1

Fusion Model For Covid-19 Diagnosis Using Chest X-Ray Images

Mohammed Hafeez M. K[1], M. Sharmila Kumari[2], Hana Fathima[3], Haleemathul Shifa Tahseen[4], Shrama K[5], Vaishnavi N Shetty[6]
Department of Computer Science and Engineering, P A College of Engineering Mangalore, *India*

****Corresponding Author***
***E-mail Id:**-hafiz123ster@gmail.com*

ABSTRACT
The COVID-19 pandemic increased at an exponential rate and faced a major problem due to the restricted accessibility of rapid test kits. The radio-imaging approach recommends that the images comprise important data related to coronaviruses. The proposed FM-HCF-DLF model is comprised of Gaussian filtering-based pre-processing, a fusion model for feature extraction and classification. The proposed FM model is a combination of handcrafted features with the help of local binary patterns (LBP), deep learning (DL) features and the convolutional neural network (CNN)-based Inception v3 technique using the Adam optimizer. The multilayer perceptron (MLP) is employed to carry out the classification process. The proposed FM-HCF-DLF model used a chest X-ray dataset for experimentation purposes. The experimental outcomes yielded superior performance with a maximum sensitivity of 93.61%, specificity of 94.56%, the precision of 94.85%, the accuracy of 94.08%, F score of 93.2% and kappa value of 93.5%.

***Keywords:**-COVID-1, Convolutional neural network, Pre- processing, Feature extraction, Fusion model Classification*

INTRODUCTION

The corona virus is around 100 nm in diameter, and 200-300 nm in length. Corona virus is a single stranded positive sense RNA enveloped by the capsid protein. The capsid contains the viral RNA as well as proteins essential for its own replication which include reverse transcriptase and integrase. Coronavirus are amongst the most pathogenic human pathogens, causing severe symptoms such as fever, vomiting and diarrhea in susceptible individuals. The coronavirus is a very small virus that you might have never heard of, but you have definitely seen its effects. Coronaviruses cause the common cold, pneumonia, and other respiratory illnesses in humans. This virus exhibits a peculiar genetic structure: it contains some genes from the herpes family of viruses and from the influenza family of viruses. Coronaviruses are one of the most widespread and important families of viruses of vertebrates and invertebrates. They are found in habitats ranging from sea water to fresh water, and from the surface to the deep sea. Their hosts range from bacteria and protozoans, to insects, fish, amphibians, reptiles, birds and mammals. The Togaviridae family contains some of the most pathogenic viruses known to infect humans. The family includes among others, the rubella virus causing congenital defects, the Chikungunya virus causing arthritic symptoms in humans and the dengue fever virus whose serious clinical picture includes hemorrhagic fever and shock. In addition, there are yellow fever, measles and mumps viruses as well as several severe acute respiratory pathogens [1].The first case was identified in December 2019

in Wuhan city, China which rapidly spread across the globe.

The number of cases has been growing since 2006 and it is predicted that in the near future we will see more outbreaks. By the end of 2010, it was reported that over 500 cases had been confirmed across Asia and Europe. Uncontrolled hemorrhaging and excruciating pain are commonly associated with Lethal Encephalitis caused by the Corona Virus. Centers for Disease Control experts believe that there are only 130 deaths caused by the Corona Virus in the United States every year. Unfortunately, due to the virus' long incubation period, many people who have been infected do not even know they have been affected.

The mosquito-borne Corona virus is the leading cause of death for children under the age of 5. It kills more than 100,000 children per year, and infects another 500,000 with painful and debilitating symptoms such as crippled limbs and cleft palates. The Corona virus is a mosquito-borne disease similar to malaria in its modes of transmission. While transmission between humans is rare, it can also be spread by human-to-human contact

This paper concentrates on an effective fusion model (FM), hand-crafted with in-depth learning features called the FM-HCFDLF model for the diagnosis and classification of COVID-19. The proposed FM-HCF-DLF model consists of Gaussian filtering (GF)-based pre- processing, FM for feature extraction and classification. FM model incorporates the fusion of handcrafted features (HCF) using local binary patterns (LBP). In contrast, deep learning features (DLF) utilize convolutional neural network (CNN) - based Inception v3 approach with Adam optimizer as a scheduler. Lastly, the multilayer perceptron (MLP)-based classification process is applied to classify COVID-19 into different sets of classes. The chest X-ray dataset was given as an input to the proposed FM-HCF-DLF and defined a superior performance of the presented model.

SOFTWARE REQUIREMENTS SPECIFICATION

A. Overall Description

Coronavirus can be transmitted through a variety of ways: oral, respiratory, etc. It is difficult to determine if someone has contracted the coronavirus because the symptoms are not always present. Patients experience influenza-like signs, such as difficulty in breathing, dry cough, tiredness, and fever. In serious cases where the person has comorbidities, i.e., affected by other diseases like blood pressure, diabetes, or heart problems, pneumonia develops rapidly resulting in acute renal failure and finally death in the worst cases. At present, the best method to determine COVID-19 is to perform a swab test and examine the biotic material collected from patients using real-time reverse transcriptase polymerase chain reaction (RT-PCR).

However, it is a challenge that the swab test is taken only for those individuals with COVID-19 symptoms. The existing COVID-19 patients without symptoms could not be recognized until they approach the hospitals. Though the disease can be diagnosed by polymerase chain reaction, COVID-19 patients who are infected with pneumonia can be diagnosed using chest X-rays and computed tomography (CT) images only. In one of the studies conducted recently, COVID-19 can be slightly identified by the human eye too. The COVID-19 transmission rate is calculated based on the volume of affected patients who are consistently diagnosed with minimum false negatives.

B. Machine learning (ML)

based applications are currently being

utilized for automatic disease diagnosis in the healthcare sector. DL, which makes it possible to develop end-to-end methods that guarantee results, is one of the most common areas of AI research. This is done without having to manually extract features from the intake data. The DL method is effective for lung segmentation, skin cancer classification, fundus image segmentation, brain disease classification, and pneumonia detection from chest X-ray images.

C. Specific Requirements

The interface should be simple and easy to understand and use. It should also be an interactive interface.

- Model construction: Model construction depends on machine learning algorithms. In this projects case, it was neural networks. At this moment Keras communicates with Tensor Flow for the construction of the model. During model compilation, it is important to write a loss function and an optimizer algorithm. The loss function shows the accuracy of each prediction made by the model.
- Model training: After model construction, it is time for model training. In this phase, the model is trained using training data and the expected output for this data. Progress is visible on the console when the script runs. In the end, it will report the final accuracy of the model.
- Model testing: During this phase, a second set of data is loaded. This data set has never been seen by the model and therefore its true accuracy will be verified. After the model training is complete, and it is understood that the model shows the right result.
- Model evaluation: Finally, the saved model can be used in the real world. This is model evaluation. This means that the model can be used to evaluate new data.

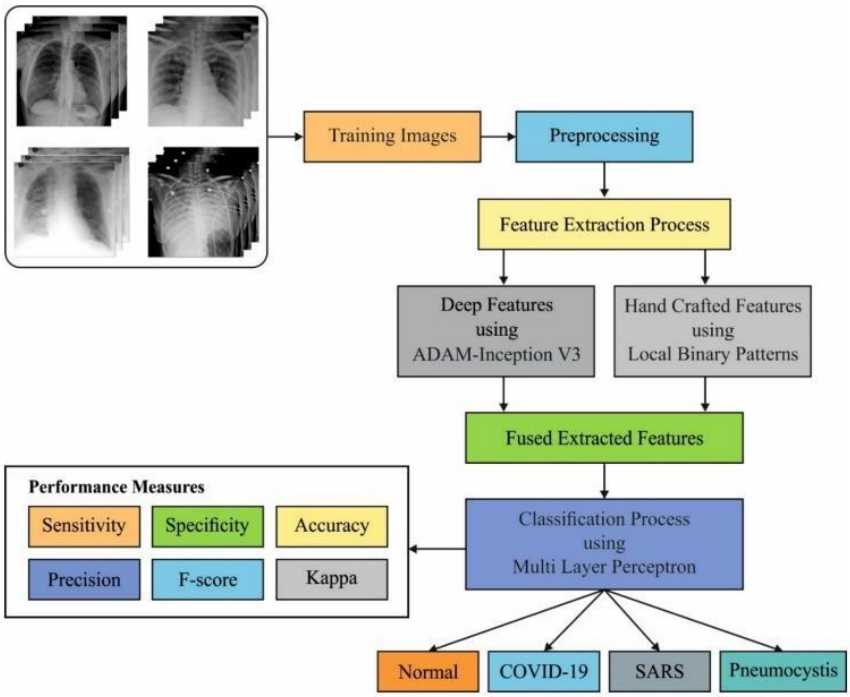

Fig.1:*-The Working of FM-HCF-DLF model*

HIGH LEVEL DESIGN

A. System design is the process of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements. One could see it as the application of systems theory to product development. The purpose of System Design is to create a

technical solution that satisfies the functional requirements of the system. The functional specification produced during system requirements analysis is transformed into a physical architecture. System components are distributed across the physical architecture, usable interfaces are designed and prototyped and technical specifications are created for the application developers, enabling them to build and test the system.

B. Design Considerations

Chest radiographs may fail to reveal pneumothorax or radiologists or interpreting physicians may fail to recognize the presence of the pneumothorax. In patients with underlying pulmonary disease, the classic visceral pleural line may be harder to detect, because the lung is hyper lucent, and little difference exists in the radiographic density between the pneumothorax and the emphysematous lung. Models with location-attention mechanisms could more accurately classify COVID-19 at chest radiography with an overall accuracy rate of 86.7 % and could be a promising supplementary diagnostic method for frontline clinical doctors. This means that there is a possibility to have remaining per cent as error in classifying the image from chest x-ray. Another drawback to chest radiographs is the image can blurred. When that happens it is difficult for the specialist to classify their illness accordingly. Apart from that some of the aspects are:

• Compatibility - The software can interoperate with other products.Extensibility - New capabilities can be added without major changes to the underlying architecture.

A. Algorithms

DETAILED DESIGN

• Fault-tolerance - The software can recover from failed components.

• Maintainability - A highly maintainable design helps to fix the bug and modifications easily and adds high modularity and extensibility to the software product.

B. Assumptions and Dependencies

This project is at its initial stage will be developed by website and then later application will be introduced. The website will contain latest technology that is latest version of Python and server language. The application is ready to use whenever necessary. It should provide the appropriate result without any error otherwise there will be huge mistake in detecting the diseases. It should provide good graphic quality. Must detect and classify Diseases based on the chest X-Ray images.

C. General Constraints

Diagnosing COVID-19 from CXR images is a complicated task for radiologists. They must identify typical patterns of the disease that are often shared with other types of viral pneumonia, which leads to errors in their diagnosis.

In order to combat COVID-19, accurate and efficient diagnostic tests must be available to detect SARS-CoV-2 and antiviral antibodies in infected individuals. SARS-CoV-2 RNA is detected early with reverse transcription-polymerase chain reaction, but isothermal nucleic acid amplification assays, including transcription-mediated amplification and CRISPR-based methods, are promising alternatives. Serological tests can detect antibodies to SARS-CoV-2, such as enzyme-linked immunosorbent assays (ELISAs) and lateral flow immunoassays. To facilitate future improvements and innovation, this paper presents a review of current COVID-19 diagnostic techniques and products.

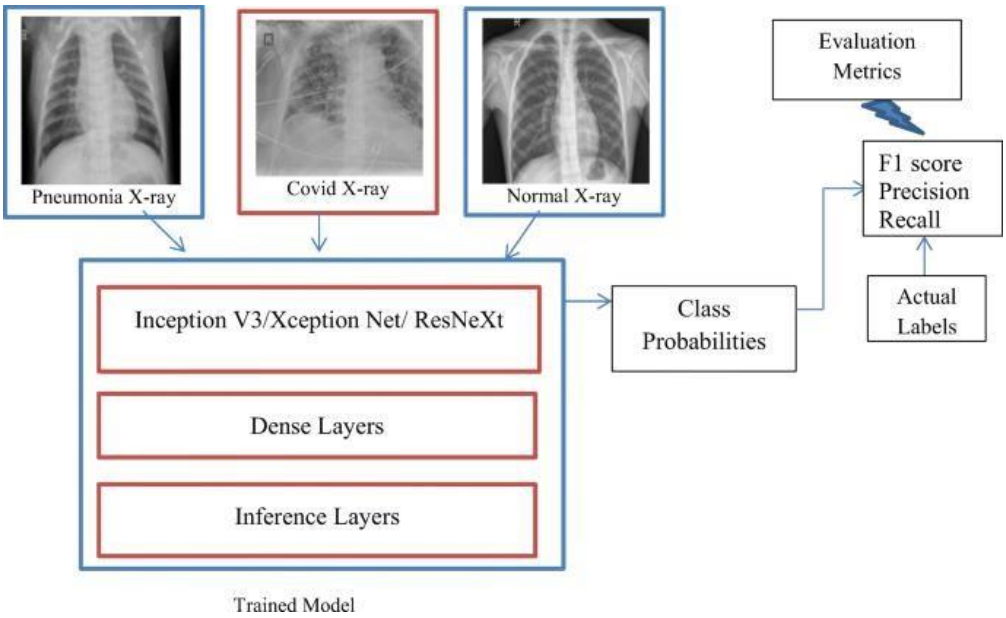

Fig.2:-*Deep learning detection and analysis of covid-19 on chest x-ray images*

 CNN Model

CNN comprises a set of convolutional layers to detect patterns that exist in the image. The merit of CNN is that it finds useful in the design of a very deep network with a few parameters for training purposes. It also reduces the time and difficulty in the training task. Besides, the CNN includes different layers namely convolutional, activation, pooling, Fully Connected (FC) layers and SoftMax (SM) layers. The basic concept lies in the image classification is horizontal or vertical edge recognition that is attained by executing a convolution operation on the input image. CNN considers a small square (or 'window') known as a filter and begins to employ it over the image. Every filter enables CNN in the identification of particular patterns in the image.

 Convolution Layer

It comprises a filter which convolves across the width and height of the input volume. Alternatively, the outcome of the convolution layer is attained by performing a dot product operation among the filter weight content and every location of the image. It leads to the generation of the two- dimensional activation map which offers the responses of that filter at each spatial location. Different variables that exist in the process are filter count, filter sizes, weight content, strides and padding.

Pooling layer

It intends to avoid over fitting and employs non-linear down-sampling on the activation map to minimize the dimensionality and complexity to increase the processing. Several variables exist in the processes are filter size and stride, whereas padding is not used in pooling. Besides, it is employed on every input channel separately; thereby the input and output channel count become identical. The two kinds of pooling such as max pooling and average pooling are defined below.

Max pooling: The fundamental nature of this layer is identical to the convolutional layer. A major variable is that it takes maximum neighbouring value from every individual location of the input image. It is carried out on every individual channel in the input.

Average pooling: In this layer, the

average of every value exists adjacent to every individual location in the input image.

IMPLEMENTATION

The implementation phase of any project development is the most important phase as it yields the final solution, which solves the problem at hand. The implementation phase involves the actual materialization of the ideas, which are expressed in the analysis document, and development in the designed phase. Implementation of any software is always preceded by important decisions regarding selection of the platform, the language used, etc. These decisions are often influenced by several factors such as the real environment in which the system works and the speed that is required, other implementation specific details, etc.

A. Platforms Used

- Jupyter Notebook

The Jupyter Notebook is a web application that allows you to create and share documents that contain live code, equations, visualizations, and explanatory text. Uses include data cleaning and transformation, numerical simulation, statistical modeling, machine learning, and much more. Jupyter Notebook is an open source web application that provides interactive computing capabilities. It has multi-language support and formats output in HTML, Latex, and many other formats. You can write code with Jupyter Notebook, create narrative text with markdown, copy/paste content with cell drag and drop, add inline images, create and run python scripts in the browser, and visualize data with Matplotlib. The Jupyter Notebook interface is available as a website or via its command line interface package 'Jupyter'. The program is used to analyze large sets of data as well as to create short bits of text with simple visualizations. Jupyter Notebook, formerly known as IPython Notebook, is an interactive computing environment for discovering, exploring, visualizing, and sharing information. Jupyter Notebook supports many programming languages including Python 2 & 3, Julia, R and Scala The interface is based on a browser: users can switch back and forth between input (code) cells and output/display cells to see the results of their code or analyses. Jupyter Notebooks can also include executable math expressions written in JavaScript which can be embedded into markdown cells.

Visual Studio Code

Visual Studio Code is a new lightweight code editor from Microsoft, built to edit all kind of files. It is open source and cross-platform, available for Windows, Linux and Mac platforms. Visual studio code is the first of its kind language agnostic editor, providing an overall unified editing experience for many languages out of the box. VSCode comes with built in support for JavaScript, TypeScript and JSON as well as support for language extensions for C++, C. Like most modern editors, it supports many programming languages out of the box, but its true strength lies in its extensibility. The VS Code team provides excellent support for additional languages (or file formats) through a variety of specialized extensions.

CONCLUSION

The FM-HCF-DLF model developed for COVID-19 diagnosis and classification involved pre-processing stage using the Gaussian technique to remove the image noise. The feature extraction process based on the FM model extracted the useful set of features from the pre-processed image. The HCF features used LBP while the DLF used CNN-based Inception v3 model with Adam optimizer to adjust the learning rate of the Inception v3 model.

REFERENCES

1. Fan Y, Zhao K, Shi Z, Zhou P (2019)

Bat coronaviruses in China. Viruses. 11:210
2. Chen Y, Liu Q, Guo D (2020) Emerging coronaviruses: genome structure, replication, and pathogenesis. J Med Virol 92:418–423
3. Setio AAA, Ciompi F, Litjens G, Gerke P, Jacobs C, van Riel SJ et al (2016) Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 35(5):1160–1169
4. Xia K, Yin H, Qian P, Jiang Y, Wang S (2019) Liver semantic segmentation algorithm based on improved deep adversarial networks in combination of weighted loss function on abdominal CT images. IEEE Access 7:96349–96358
5. Pezeshk A, Hamidian S, Petrick N, Sahiner B (2019) 3-D convolutional neural networks for automatic detection of pulmonary nodules in chest CT. IEEE J Biomed Health Inform 23(5):2080–2090
6. Prakasa E (2016) Texture feature extraction by applying local binary pattern (LBP). INKOM J 9(2):45–48
7. Nardelli P, Jimenez-Carretero D, Bermejo-Pelaez D, Washko GR, Rahaghi FN, Ledesma-Carbayo MJ et al (2018) Pulmonary artery–vein classification in CT images using deep learning. IEEE Trans Med Imaging 37(11):2428–2440.