Decision Support System Based on Deep Learning with Big Data for Pneumonia Prediction

Dr. Anupa Sinha[1], Yalakala Dinesh Kumar[2]

[1]*Assistant Professor, Department of CS & IT, Kalinga University, Raipur, India.*
[2]*Research Scholar, Department of CS & IT, Kalinga University, Raipur, India.*

KEYWORDS

Health, public health, pneumonia, Decision support system

ABSTRACT

A bacterial, viral, or fungal infection of the lungs causes pneumonia in humans. It is the primary cause of global decline. Around 27% of deaths worldwide are attributable to this illness. A World Health Organisation report states that this serious infectious disease claims the lives of almost 2.5 million people annually. Early detection of chest illness, including pneumonia, can reduce the likelihood that the condition will worsen by allowing for the proper course of treatment. Accurate CT scans and CXR (chest X-ray) reports are renowned for being exact methods of diagnosis and prediction. This approach is trustworthy and non-invasive, and it can be used as a screening tool when a disease diagnosis is necessary. Nevertheless, research has demonstrated that even highly trained radiologists regularly fail to correctly identify pneumonia disease in its early stages. The severity of the sickness and the death rate increase as a result of diagnostic delays. Machine learning has affected every aspect of human life in the twenty-first century. It has a significant effect on the medical field in terms of early diagnosis of serious illnesses. Therefore, an intelligent health care system based on artificial intelligence (AI) was presented in this work for the early and accurate diagnosis of pneumonia disease.

1. Introduction

Increased mortality is caused by the fatal illness pneumonia. Most of the time, pneumonia does not exhibit any symptoms at first. Symptoms started to appear as it became worse. Physical detection techniques are expensive and time-consuming. Decision support systems and deep learning techniques can be used to diagnose pneumonia. This may lessen the possibility of human error in pneumonia diagnosis [1]. A decision support system can assist in the diagnosis process before a condition worsens. Timely diagnosis is crucial in preventing lung disorders like pneumonia and perhaps saving lives [12]. Decision Support System (DSS) is becoming more and more common in the healthcare sector because it helps physicians and patients alike. The resources available in underdeveloped nations are scarce [2]. Therefore, it's important to use the resources wisely. The lack of physicians and medical personnel is a difficult issue in the twenty-first century. Access to a decision support system can facilitate doctors' job by enabling them to make optimal decisions in a quick, accurate, and early manner. This will enable medical professionals to treat individuals' illnesses more precisely. Additionally, it might help patients avoid the need for additional costly testing, which would lead to more economical care [11]. Decisions in the clinical domain are made by DSS using information gathered from multiple sources. It can help doctors diagnose patients more rapidly, recommend preventive measures to patients, and administer individualized care. It can be applied to early disease prediction. This could suggest the best course of action for a patient. DSS can be used to identify risk factors for disease [3]. It is possible to identify an aberrant state in a patient and produce alarm messages accordingly. It can provide the patient with advice on lifestyle changes to prevent contracting the illness. Data can be obtained for decision-making from a number of sources. Information obtained from remote sensors, X-ray, CT, and other medical test results, doctor's notes, and information from websites like Kaggle and UCI are all included in this. DSS has the ability to find hidden patterns in data and form conclusions based on those patterns. Modern technology has advanced to the point that DSS is now a helpful tool in a wide range of industries. The use of decision support systems, or DSS, is essential to making wise decisions. Deep Learning has had a significant impact on the healthcare industry, and numerous DSS have been developed for the purpose of illness diagnosis.

2. Literature Review

Although pneumonia is a fatal illness, its severity can be decreased with prompt diagnosis. Decision support systems based on machine learning can help in pneumonia prediction. Numerous efforts have been made and will continue to be made in this regard by researchers [22]. An extensive overview of

the several Decision Support Systems created for the diagnosis of pneumonia is given in this chapter. Performance metrics and other techniques of validation were also provided. Additionally included are the parameters and validation techniques used to assess these systems' performance. To validate their algorithms, researchers have used a number of online databases related to pneumonia sickness.

Several techniques for improving the accuracy of medical photographs were presented in [4]. Various preprocessing techniques are available to eliminate unnecessary and unsuitable data. Better outcomes can be achieved by starting the image processing process with these pretreatment techniques [6]. In order to improve the quality of the photographs, the authors of this study compared various variance reduction methods. A secure system for the BigData healthcare lifecycle was given in [5]. The goal of big data is to gather, store, preprocess, and convert massive amounts of data into information that can be used to techniques and technologies. When big data techniques are applied properly, they can be utilised to reduce medical expenditures and improve disease detection. The primary areas of interest for this study include security, access control, encryption, and authentication. Utilise machine learning algorithms and the BigData idea in [14] to effectively anticipate chronic diseases. The data gathered from the Central China Hospital served as the foundation for this investigation. The latent factor model is utilised to reconstruct the missing data in order to solve the incomplete data issue. Using both structured and unstructured data, the authors demonstrated the CNN-MDRP algorithm [9], which can reach a 94.8% convergence rate. R. [7] spoke about the use of BigData for illness prediction in the medical industry. Using the Naive Bayes Technique, the BigData Predictive analysis model (BPA-NB) was employed by the authors to predict diseases. For the goal of classification, the probabilistic technique is employed. BigData in particular is a good fit for the Naive Bayes techniques because of its enormous data sets. The Naive Bayes algorithm was trained using data from the UCI Repository. The outcomes can then be estimated using the data that has been collected. The method used here demonstrates that BPA-NB approaches have a 97.12% prediction accuracy for illnesses.

3. Methodology

Structured learning, machine learning, and hierarchical learning are all combined to create deep learning. It is based on a set of predefined algorithms that use several structures nonlinear transformations and compound formats to attempt to alter the intelligence of various data levels utilising different processing layers. This learning approach may be considered as an entire collection of many machine learning processes that rely on learning the data presentation method. The image can be displayed in a variety of ways, including areas of a specific size, a set of edges, and values per pixel intensity. This technique's most significant ability is its capacity to replace some variants with efficient methods for extracting structural features and learning different features. Owing to technical progress, these methods are important in developing a tool for diagnosing diseases. The main driving force behind this strategy is the extreme scarcity of health care services in developing nations like India. Deep learning can assist in making more precise and accurate disease predictions. [8].

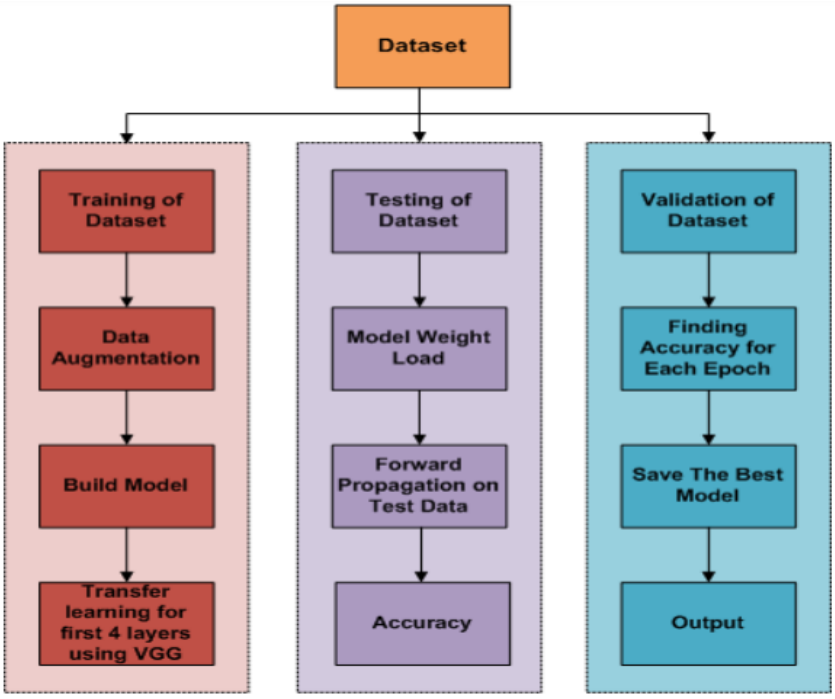

Figure 1: overall proposed framework

Data scientists and data learners are mostly in charge of managing big data. In the context of healthcare, "big data" refers to the collection, examination, analysis, and utilisation of patient data databases that are too large or complicated for conventional data processing techniques to comprehend. It can be defined as the vast quantity of data that is produced daily from various stages. Large patient records are gathered using big data, which benefits the healthcare sector by preserving hospital performance. Deep learning is a particularly helpful technique for big data analysis because its primary goal is to mine massive amounts of data for patterns and relevant features from composite data without the need for human intervention. From CXR data, diseases can be identified using the artificial intelligence-based convolution neural network (CNN) VGG-16. The several approaches used by researchers to diagnose medical conditions like pneumonia are covered in this section [13].

The purpose of this submission was to suggest the Advanced VGG-16 algorithm. From the collection of X-ray pictures, a convolution neural network based on transfer learning and fine-tuning advanced VGG-16 has been analysed to diagnose pneumonia. For improved outcomes, techniques like data augmentation and fine tuning are applied. The cropped sequential residual model and the baseline customised model were used to compare the outcomes of the advanced VGG-16 model [10].

4. Results and discussion

The findings in terms of accuracy, specificity, precision, recall, and F-measure were obtained following the deployment of the upgraded VGG-16 model. The relationship between Loss and Epoch is depicted in Figure 2's plotted graph above. The loss curve is diminishing as the number of epochs increases.

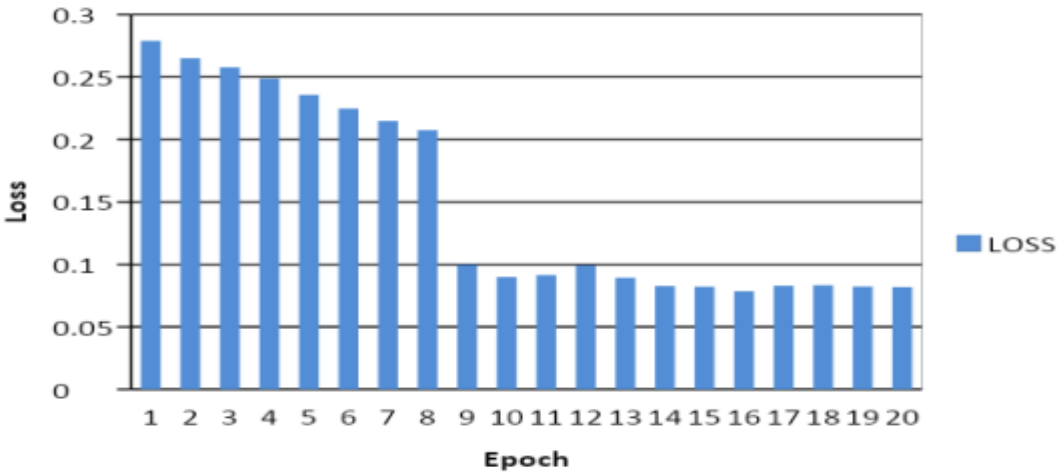

Figure 2: Loss vs. Epoch.

The curve connecting accuracy and epoch is depicted in Figure 3 of the plotted graph above. The accuracy curve increases as the number of epochs increases. The graph below, which is shown according to Figure 4, displays the relationship between loss, accuracy as measured, and epoch. The loss curve is reducing as the number of epochs increases, while the accuracy curve is concurrently increasing.

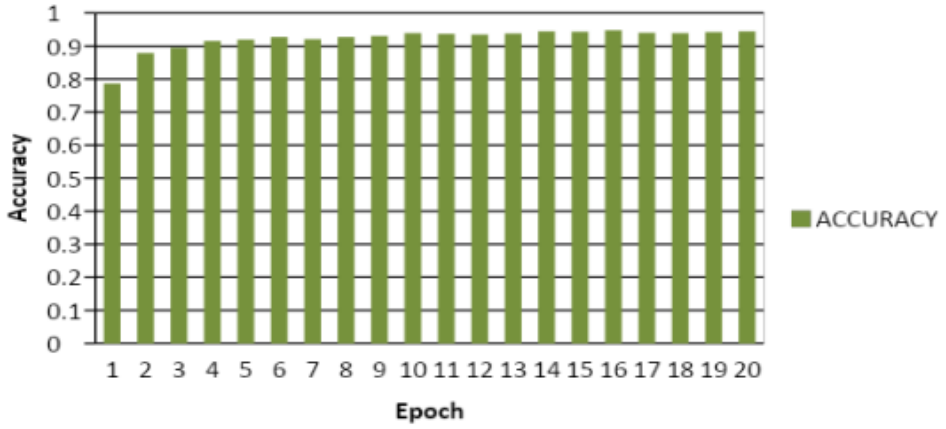

Figure 3: Accuracy vs. Epoch.

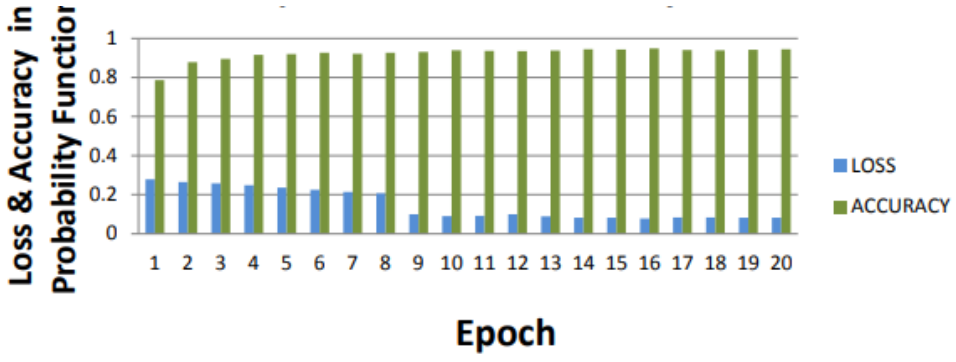

Figure 4: Epochs Vs. Loss and Accuracy in Probability Function

The results of the advanced VGG16 models were found to be superior to those of the other models, as indicated in Table 3, and their parameters are further displayed in Figure 4. In the following step, we compared the performance of the Advanced VGG16 Model with the other models, such as the Baseline Customised VGG16 model and the cropped sequential residual model, in terms of various performance parameters.

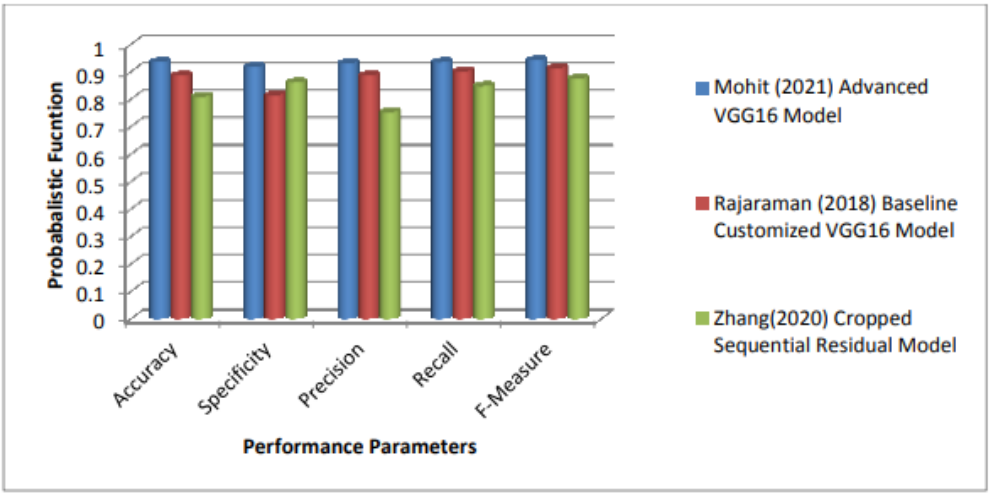

Figure 4: Comparison with Existing Methods

The primary goal of the authors' study is to create an enhanced diagnostic system for the prediction of pneumonia from the dataset of X-ray pictures. They do this by concentrating on the advanced VGG16-based model and some of its other characteristics, such as augmentation, shearing, and zooming. It is possible to obtain various parameters by utilising this model. Three different kind of values are really used by our model: test, train, and validate. We may quickly divide the data into two groups, such as normal (0) and pneumonia (1), by giving the input values.

5. Conclusion and future scope

Developing a smart healthcare system that can be utilised to diagnose pneumonia using various technologies is the main goal of the proposed dissertation. Here, the focus is on improving and analysing a system that uses clinical criteria to identify pneumonia as well as other pathological conditions. Numerous deaths from this acute pneumonia occur each year. If this illness is not identified in a timely manner, it may be fatal. Timeliness and accuracy are crucial factors.This is a crucial component in the diagnosing process.This terminology yields results that are accurate and encouraging. Disease detection is a serious challenge in poor nations like India due to a lack of medical resources. In this study, a sophisticated VGG-16 model for identifying pneumonia in medical photographs was examined. For improved outcomes, techniques like data augmentation and fine tuning are applied. The many pre-established performance criteria were used as the foundation for the outcomes that were attained. The cropped sequential residual model and the baseline customised model were used to compare the outcomes of the advanced VGG-16 model.

Reference

[1] Sandeep Kumar, E., and Pappu Satya Jayadev. "Deep learning for clinical decision support systems: a review from the panorama of smart healthcare." *Deep learning techniques for biomedical and health informatics* (2020): 79-99.

[2] Singh, Swapnil, Ameyaa Biwalkar, and Vidhi Vazirani. "Clinical Decision Support Systems and Computational Intelligence for Healthcare Industries." In Knowledge Modelling and Big Data Analytics in Healthcare, pp. 37-63. CRC Press, 2021.

[3] Alamer, L., Alqahtani, I. M., & Shadadi, E. (2023). Intelligent Health Risk and Disease Prediction Using Optimized Naive Bayes Classifier. Journal of Internet Services and Information Security, 13(1), 01-10.

[4] Kindle, Ryan D., Omar Badawi, Leo Anthony Celi, and Shawn Sturland. "Intensive care unit telemedicine in the era of big data, artificial intelligence, and computer clinical decision support systems." Critical care clinics 35, no. 3 (2019): 483-495.

[5] Farahani, Ali, Toktam Khatibi, Hossein Sarmadian, and Azam Boskabadi. "Proposing a two-step decision support system for differential diagnosis of tuberculosis from pneumonia." Sustainable Operations and Computers 3 (2022): 303-316.

[6] Ajwad, A.A., Ahmed, A.A., Kamal, M., Jaleel, R.A., & Mahmood, M.B. (2023). Improved Secure IoTs-Based Visual Computing with Image Processing and Artificial Intelligence Techniques for Accurate Predicting of Novel COVID. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 14(1), 1-14.

[7] Karthikeyan, Divydharshini, Aparna S. Varde, and Weitian Wang. "Transfer learning for decision support in Covid-19 detection from a few images in big data." In 2020 IEEE International Conference on Big Data (Big Data), pp. 4873-4881. IEEE, 2020.

[8] Qjidaa, M., A. Ben-Fares, Y. Mechbal, H. Amakdouf, M. Maaroufi, B. Alami, and H. Qjidaa. "Development of a clinical decision support system for the early detection of COVID-19 using deep learning based on chest radiographic images." In 2020 International Conference on Intelligent Systems and Computer Vision (ISCV), pp. 1-6. IEEE, 2020.

[9] Kutlu, Y., & Camgözlü, Y. (2021). Detection of coronavirus disease (COVID-19) from X-ray images using deep convolutional neural networks. Natural and Engineering Sciences, 6(1), 60-74.

[10] Jakhar, Karan, and Nishtha Hooda. "Big data deep learning framework using keras: A case study of pneumonia prediction." In 2018 4th International Conference on computing communication and automation (ICCCA), pp. 1-5. IEEE, 2018.

[11] Yau, Ashiru Anees, Saiffuddin Kamfut Sani, Aparna Datta, and . 0. Potential Of Curcumin Loaded Nanoparticles In Antimicrobial Photodyanamic Therapy. doi:10.31838/ijprt/11.02.07

[12] do Amaral, Jorge Luis Machado, and Pedro Lopes de Melo. "Clinical decision support systems to improve the diagnosis and management of respiratory diseases." In Artificial intelligence in precision health, pp. 359-391. Academic Press, 2020.

[13] Ramakrishnan, J., Ravi Sankar, G., & Thavamani, K. (2019). Publication Growth and Research in India on Lung Cancer Literature: A Bibliometric Study. Indian Journal of Information Sources and Services, 9(S1), 44–47.

[14] Mayya, Veena, K. Karthik, Kamath S. Sowmya, Krishnananda Karadka, and Jayakumar Jeganathan. "COVIDDX: AI-based Clinical Decision Support System for Learning COVID-19 Disease Representations from Multimodal Patient Data." In HEALTHINF, pp. 659-666. 2021.

[15] Wu, Guangyao, Pei Yang, Yuanliang Xie, Henry C. Woodruff, Xiangang Rao, Julien Guiot, Anne-Noelle Frix et al. "Development of a clinical decision support system for severity risk prediction and triage of COVID-19 patients at hospital admission: an international multicentre study." *European Respiratory Journal* 56, no. 2 (2020).