IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 12, Issue 9, September 2023

DOI: 10.17148/IJARCCE.2023.12922

The Impact of Machine Learning in COVID-19 Detection and Diagnosis

Dr. Santosh Jagtap[1]

Assistant Professor, Department of Computer Science, Prof. Ramkrishna More College, Pune, India[1]

Abstract: Machine learning models for COVID-19 detection using chest X-ray images are important because they are fast, accurate, non-invasive, and accessible. They can be used to identify patients who are at high risk of complications, monitor the progression of the disease, and develop new diagnostic and treatment strategies. The ultimate goal of the research is to develop an accurate, reliable, and cost-effective automated diagnostic tool for COVID-19 detection using CXR images, which can help to reduce the spread of the disease and improve patient outcomes. In this study, the researcher invented a new system for COVID-19 detection based on image processing with chest X-ray images. The primary focus of the experiment is to detect various diseases using chest X-ray images.

Keywords: KNN, RF, NN

I. INTRODUCTION

Machine learning models for COVID-19 detection using chest X-ray images are important because they could help to reduce the spread of the virus and improve patient outcomes. These models can be used to quickly and accurately identify patients who are infected with the virus, even at an early stage, and to monitor the progression of the disease in patients who have already been diagnosed. Machine learning models could also be used to develop new diagnostic and treatment strategies.

Recent advances in image processing, machine learning, transfer learning, deep learning, and ensemble learning have revolutionized the field of medical diagnosis. Chest X-rays are a common imaging modality for screening and diagnosing lung diseases, and machine learning techniques such as neural networks, support vector machines, K-nearest neighbors (KNN), random forests (RF), convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks are being used for automatic disease detection. Early detection and evaluation of disease progression are critical for timely medical intervention.

Medical images provide crucial information for experts to make informed decisions. Numerous studies have investigated the use of artificial intelligence methods for automatic disease classification using chest X-rays.

The combination of machine learning with deep learning algorithms enables accurate and efficient analysis of large datasets, facilitating rapid COVID-19 detection from chest X-ray images. These algorithms are trained on a large number of images, enabling them to recognize patterns and features indicative of COVID-19 infection. This provides clinicians with a valuable tool for early detection of the disease and timely treatment, which is critical in controlling the spread of the virus.

II. REVIEW OF LITERATURE

Lamin Saidy et al. propose an efficient deep-learning method for segmenting lungs in chest X-ray images. They apply manual reference segmentation, model training, and testing, along with contour operations to enhance accuracy. The study demonstrates practical viability for real-world medical applications, providing precise lung area measurements and aiding in diverse medical diagnoses.. [1]

Tatiana Chakravorti et al. (2020) in their paper, "Detection and Classification of COVID-19 using Convolutional Neural Network from Chest X-ray Images," employ both multi-class and binary classification methods to achieve accurate COVID-19 detection. The study focuses on classifying images into three categories: COVID-19, Normal, and pneumonia. Notably, the precision and recall values for all categories in this research are exceptionally high, making it a noteworthy aspect of their findings.[2]

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁑ Peer-reviewed / Refereed journal ⁑ Vol. 12, Issue 9, September 2023
DOI: 10.17148/IJARCCE.2023.12922

In this regard, Extreme learning machine (ELM) have been compared with the proposed classifier to have a vast comparative study. In this paper author proposed a deep transfer learning classifier where a tensor flow-based CNN has been used for the classification of COVID 19. Authors conclude that accuracy and sensitivity achieved by proposed model is better than binary classification and other established algorithms.[2]

Researchers have extensively reviewed prior research and related work in the domain of image processing for medical diagnosis, using various types of medical images such as X-rays, CT scans, ultrasounds, MRIs, etc., while leveraging a range of machine learning techniques, including deep learning and transfer learning. This chapter provides a comprehensive review of studies related to both conventional and contemporary methods for disease diagnosis.

As this research focuses on the utilization of X-ray images for detecting Covid-19, a detailed survey of existing literature in this area is briefly presented. The researcher has thoroughly reviewed numerous articles, papers, e-books, journals, magazines, and dissertations related to CNNs, deep learning, machine learning, and transfer learning to gain a comprehensive understanding of the background work in previous research and to identify research gaps for the development of new methods aimed at achieving more effective results.

To enhance the accuracy of Covid-19 detection using chest X-ray (CXR) images, the researcher has undertaken the customization of pre-trained models. Additionally, an innovative system has been proposed, incorporating ensemble algorithms and a novel model designed to enhance accuracy in Covid-19 detection.

III. RESEARCH METHODOLOGY

The following steps involved in image processing for the purpose of detecting COVID-19.

Step 1: Acquisition of Images:
Image acquisition marks the initial phase in the image processing workflow. For COVID-19 detection, commonly employed imaging techniques include CT scans, which offer a comprehensive 3D view of the lungs, and X-rays, which provide a 2D image. These images are captured using specialized equipment and are typically saved in the DICOM format, a standard for medical images.

Step 2: Pre-processing:
Following image acquisition, the subsequent stage involves pre-processing. Here, acquired images undergo pre-processing to eliminate any noise or artifacts. Noise can originate from various sources like equipment or movement artifacts, potentially impacting the accuracy of image analysis. Techniques like filtering and segmentation are employed to enhance image quality and eliminate unwanted noise.

Step 3: Feature Extraction:
In the third step of the image processing workflow, feature extraction takes place. This involves the extraction of relevant features from the pre-processed images. Techniques like texture analysis, shape analysis, and statistical analysis are utilized for this purpose. Extracted features play a vital role in distinguishing between COVID-19 positive and negative cases.

Step 4: Classification:
The subsequent step is classification, wherein the extracted features are employed to identify the images as COVID-19 positive or not. This task is accomplished using a combination of machine learning algorithms like random forests, artificial neural networks (ANNs), and support vector machines (SVMs). Enhancing the model's accuracy involves utilizing a substantial dataset for training and testing.

Step 5: Post-processing:
The final step is post-processing, where the results obtained from the classification are refined to eliminate any false positives or false negatives. This step is crucial for ensuring the highest possible accuracy in COVID-19 detection. Techniques such as thresholding and morphological operations are employed to rectify false positives and false negatives.

A. ANALYSIS AND DISCUSSION

In this study, the researcher conducted an investigation into the identification of Covid-19 through the analysis of CXR images. The study encompasses various emotional classes, namely Covid-19, Normal, Lung Opacity, and Viral Pneumonia. Furthermore, the researcher underscored the substantial advantages of automated Covid-19 detection for both industry and society.

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⋇ Peer-reviewed / Refereed journal ⋇ Vol. 12, Issue 9, September 2023

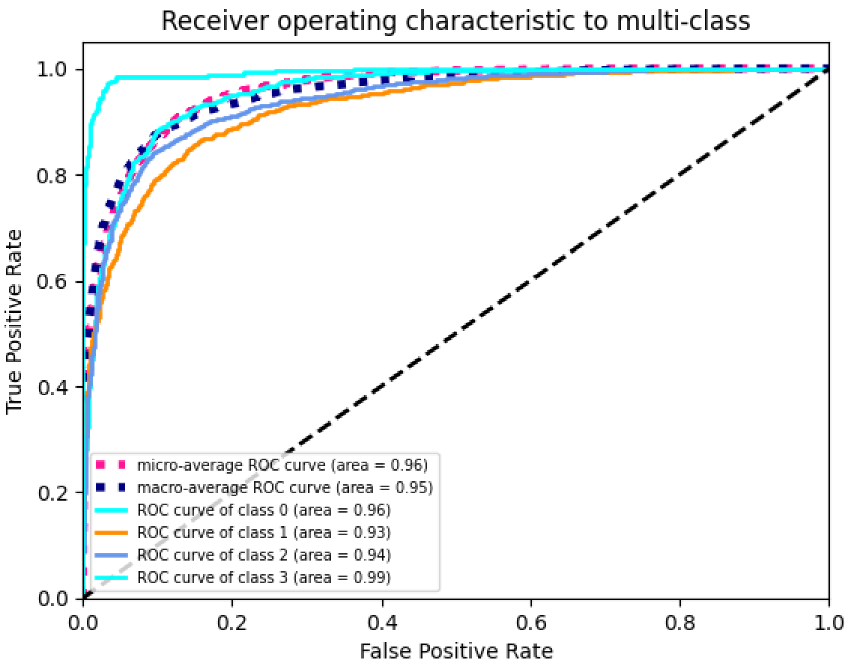
DOI: 10.17148/IJARCCE.2023.12922

The proposed framework for automatic Covid-19 detection relies on a combination of machine learning (ML) and deep learning (DL) classification algorithms. The researcher introduced four distinct types of image-based activity detection systems leveraging ensemble learning and transfer learning techniques.

ROC results of proposed models

The researcher analyzes proposed models using ROC curves. ROC curves with multiclass as Covid-19, Lung Opacity, Viral Pneumonia, and Normal lungs X-ray with micro and macro averages.

ROC of NN+RF



When it comes to differentiating between COVID-19 and non-COVID-19 situations, ensemble learning that uses both Random Forest (RF) and Neural Networks (NN) classifiers performs remarkably well. With an AUC value of 0.96, the ROC curve for COVID-19 indicates that the ensemble model has a great capacity to correctly classify occurrences of COVID-19. This shows that the ensemble model has strong discriminative abilities.

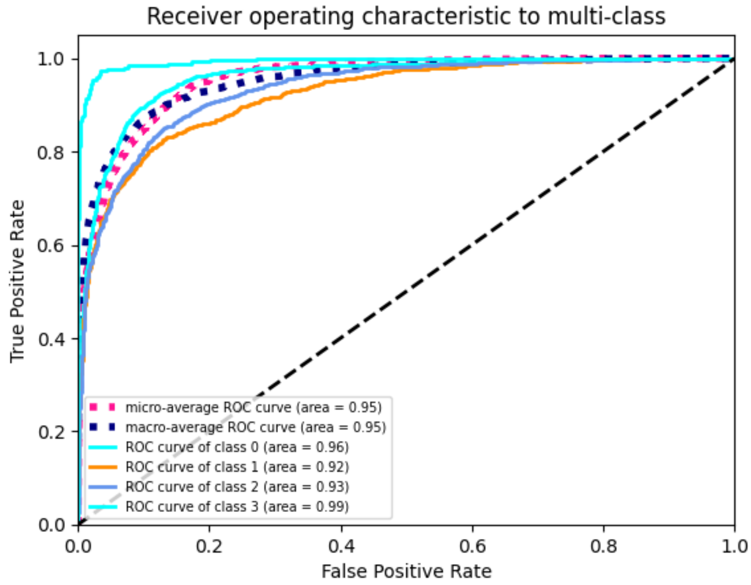
Additionally, the ensemble model performs well in separating cases of lung opacity from those without lung opacity. With an AUC value of 0.93, the ROC curve for the Lung Opacity class shows that the ensemble model is successful in capturing the specific patterns linked to lung opacity. This suggests that the ensemble model is proficient in correctly differentiating lung opacity events.

The ensemble model, made up of RF and NN classifiers, excels in properly differentiating between normal cases and abnormal cases when applied to the Normal class. The ROC curve for the Normal class displays the ensemble model's great skills in properly categorizing normal examples, with an AUC value of 0.94.

Additionally, the ensemble model performs very well in distinguishing between viral and non-viral pneumonia patients. The impressive AUC value of 0.99 displayed by the ROC curve for viral pneumonitis highlights the ensemble model's exceptional discriminatory power for this particular class.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁘ Peer-reviewed / Refereed journal ⁘ Vol. 12, Issue 9, September 2023
DOI: 10.17148/IJARCCE.2023.12922

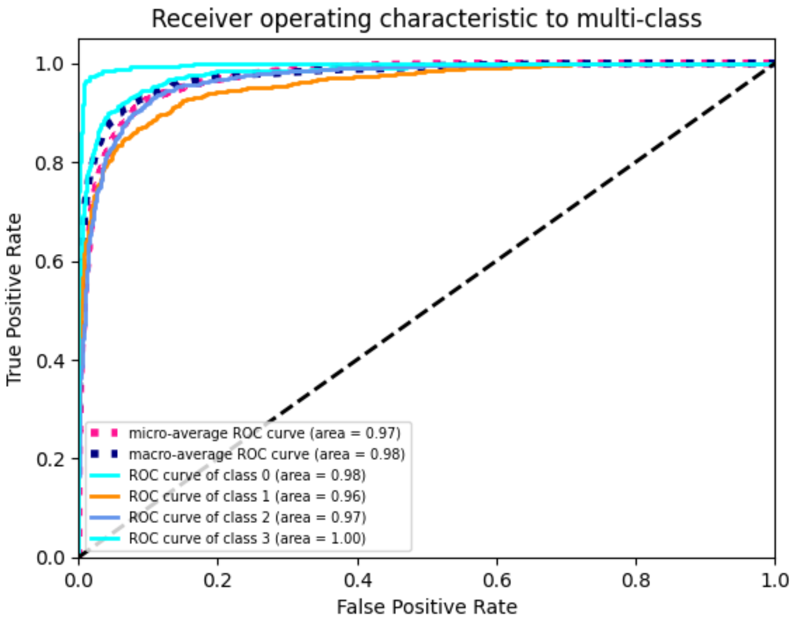
ROC of SVM+NN+RF



The ensemble model, which combines SVM, RF, and NN classifiers, performs exceptionally well at correctly categorizing COVID-19 instances. The ROC curve for COVID-19 shows that the ensemble model has a strong ability to discriminate between COVID-19 occurrences and non-COVID-19 cases, with an AUC value of 0.96. The ensemble model performs exceptionally well in correctly classifying occurrences of lung opacity, as seen by the ROC curve for the Lung Opacity class, which has an AUC value of 0.92.

The ensemble model made up of SVM, RF, and NN classifiers does a good job of reliably differentiating normal instances from abnormal ones for the Normal class. The ensemble model's capacity to accurately classify instances of the normal state is demonstrated by the ROC curve for the Normal class, which has an AUC value of 0.93. When distinguishing between cases of viral pneumonia and non-viral pneumonia, the ensemble model performs exceptionally well. The ROC curve for viral pneumonitis demonstrates the ensemble model's remarkable discriminating ability for this class with an extraordinary AUC value of 0.99.

ROC of CNN + DenseNet121(10epochs)



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

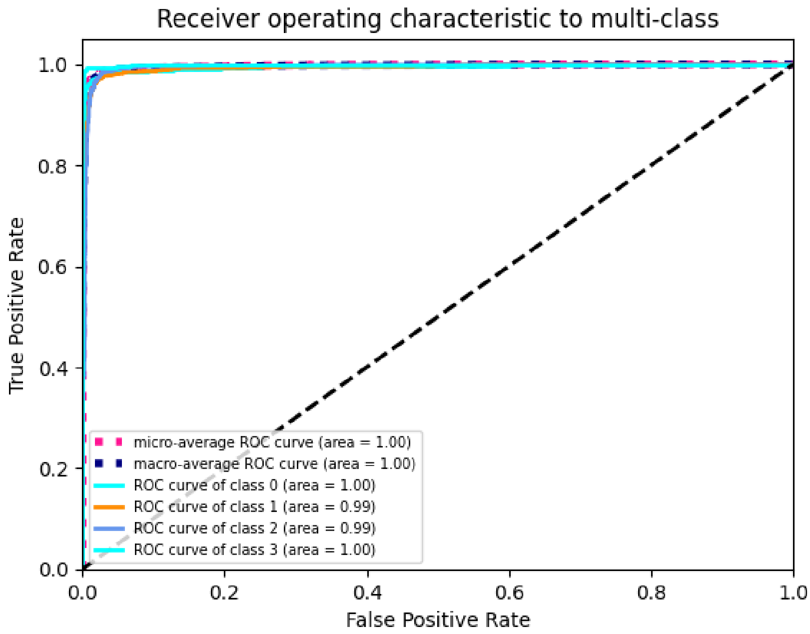
International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 12, Issue 9, September 2023

DOI: 10.17148/IJARCCE.2023.12922

The ROC curves for each class demonstrate the strong performance of the transfer learning model, which combines CNN and DenseNet121 architecture and was trained for 10 epochs. The COVID-19 class has an AUC value of 0.98, indicating that the model can reliably distinguish between COVID-19 cases and non-COVID-19 cases with a good balance between false positive and true positive rates. Similarly, the Lung Opacity class has an AUC value of 0.96, indicating that the model performs well in distinguishing between cases with lung opacity and those without. The Normal class has an AUC value of 0.97, indicating that the model can distinguish between typical and atypical instances with a high true positive rate and a comparatively low false positive rate. The Viral Pneumonia class exhibits exceptional performance, with an AUC value of 1. This shows that the transfer learning model performs exceptionally well in distinguishing cases of viral pneumonia from those of non-viral pneumonia, with a very low false positive rate and a high true positive rate. Overall, these findings demonstrate the strong discriminatory power and accuracy of the transfer learning model in categorizing various classes. These results highlight the effectiveness of transfer learning in utilizing pre-trained models and optimizing them for particular categorization tasks.

ROC of CC+DenseNet121(25epochs)



With an AUC value of 1, the ROC curve for the COVID-19 class exhibits remarkable performance. This shows that the CNN and DenseNet121 transfer learning model achieves perfect discrimination when separating COVID-19 instances from non-COVID-19 cases. Indicating a low false positive rate and a high true positive rate for COVID-19 classification, the curve is positioned towards the top-left corner.

The ROC curve performs exceptionally well for the Lung Opacity class, with an AUC value of 0.99. This shows that the transfer learning model excels at reliably distinguishing between cases of lung opacity and patients without lung opacity. The lung opacity classification curve is positioned towards the top-left corner, indicating a high true positive rate and a relatively low false positive rate. The lung opacity classification curve is positioned towards the top-left corner, indicating a high true positive rate and a relatively low false positive rate.

IV. CONCLUSION

The experiment revolves around the utilization of Covid-19 and CXR datasets for training and testing the proposed model. The results are rigorously scrutinized to gauge the model's efficacy in disease detection.

For a comprehensive analysis, the outcomes derived from the proposed model are presented and thoroughly discussed using a confusion matrix. This matrix provides valuable insights into the model's accuracy and its proficiency in accurately categorizing various diseases based on chest X-ray images. Additionally, the chapter encompasses a comparative evaluation of performance metrics across different models, offering a nuanced grasp of their respective strengths and limitations.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⧖ Peer-reviewed / Refereed journal ⧖ Vol. 12, Issue 9, September 2023

DOI: 10.17148/IJARCCE.2023.12922

REFERENCES

[1]. Lamin Saidy, Chien-Cheng Lee (2018) in," Chest X-Ray Image Segmentation Using Encoder-Decoder Convolutional Network", "IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)", ISBN:978-1-5386-6302-8.

[2]. Tatiana Chakravorti , Vinay Kumar Addala, J. Shivam Verma(2021)," Detection and Classification of COVID 19 using Convolutional Neural Network from Chest X-ray Images", "2021 6th International Conference for Convergence in Technology (I2CT)", ISBN:978-1-7281-8876-8

[3]. Ankan Ghosh Dastider, Mohseu Rashid Subah, Farhan Sadik, Tanvir Mahmud, and Shaikh Anowarul Fattah(2020)," ResCovNet: A Deep Learning-Based Architecture For COVID-19 Detection From Chest CT Scan Images", "IEEE REGION 10 CONFERENCE (TENCON)", ISBN:978-1-7281-8455-5.

[4]. Alaa A. Saif, Borhan M. N. Radman, Ahmed Y. A. Saeed, Mohammed Hashem Almourish (2021), " Covid-19 Diagnosis Based on CT Images Using Pre-Trained Models", " International Conference of Technology, Science and Administration (ICTSA)", 978-1-6654-2862-0

[5]. Shubham Chaudhary, Sadbhawna, Vinit Jakhetiya, Badri N Subudhi, Ujjwal Baid , Sharath Chandra Guntuku, " Detecting Covid-19 and Community Acquired Pneumonia Using Chest CT Scan Images With Deep Learning", "ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)", ISBN:978-1-7281-7605-5

[6]. Shayan Hassantabar, Mohsen Ahmadi, Abbas Sharifi, "Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches", Chaos, Solitons & Fractals, Volume 140,2020,110170, ISSN 0960-0779

[7]. Subhaluxmi Sahoo; Sony Snigdha Sahoo, "A new COVID-19 medical image steganography based on dual encrypted data insertion into minimum mean intensity window of LSB of X-ray scans". ISBN:978-1-7281-6917-0

[8]. R.Helen, Dr.N.Kamaraj, Dr.K.Selvi, V.Raja Raman (2011) " Segmentation of pulmonary parenchyma in CT lung images based on 2D Otsu optimized by PSO", " International Conference on Emerging Trends in Electrical and Computer Technology", ISBN No. 978-1-4244-7926-9

[9]. Shintaro Suzuki, Xiaoyong Zhang, Noriyasu Homma, Kei Ichiji, Norihiro Sugita, Yusuke Kawasumi, Tadashi Ishibashi, Makoto Yoshizawa (2016), " Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis", " 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE) ", ISBN:978-1-5090-0937-4

[10]. Ahmed Iqbal Pritom, Md. Ahadur Rahman Munshi, Shahed Anzarus Sabab, Shihabuzzaman Shihab(2016), "Predicting breast cancer recurrence using effective classification and feature selection technique", " 9th International Conference on Computer and Information Technology (ICCIT)", ISBN:978-1-5090-4091-9.

[11]. Yen-Wei Chen, Rui Xu, Yi Wang, Tiantian Liu, Xinchen Ye, Lin Lin, Shoji Kidok and Noriyuki Tomiyama (2020), "BG-Net: Boundary-Guided Network for Lung Segmentation on Clinical CT Images", "25th International Conference on Pattern Recognition (ICPR)", ISBN:978-1-7281-8808-9