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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141166

Prediction of COVID-19 Severity by Applying Machine and Deep Learning Techniques

Vishakha Aggarwal[1], Dr. Vikas Shrivastava[2]

M.Tech CSE, Meerut Institute of Engineering and Technology, Meerut[1]

Professor, Dept of CSE, Meerut Institute of Engineering and Technology, Meerut[2]

Abstract: Due to the COVID-19 outbreak, greater and more dependable tools were required to predict the severity of the disease and help specialists in their decision-making. Conventional approaches might be ineffective and time-consuming in the event of handling large and complex patient data. This paper presents a computer-aided diagnostic model, which is based on machine learning and the application of deep learning to predict the correct outcome. The model uses UNET to process medical images to segment them, and CNN/ResNet50 to classify the chest X-ray or CT scan. In order to enhance accuracy, a hybrid approach is formed by using optimized features with a Random Forest classifier. The system is programmed and developed to have a simple Graphical User Interface (GUI) that allows an individual to upload medical images and get an automated prediction with visualized output. Accuracy, precision, recall, and F1-score are used to compare performance, and the findings indicate that the proposed model is better than the current methods that can be used to offer an effective and convenient means of detecting and predicting the severity of COVID-19 at an early stage.

Keywords: COVID-19 Prediction; Machine learning; Deep learning; UNET; CNN; ResNet50; Random Forest; Image Segmentation; Medical Diagnosis; Computer-Aided Detection.

I. INTRDUCTION

The severity of COVID-19 has been predicted using machine learning (ML) and deep learning (DL) methods actively. Keerthi and Naga Raju [1] demonstrated that the classical ML models are less effective at processing complex patient data as compared to the DL models, yet interpretability remains an issue. Likewise, Sayed et al. [2] utilized the images of chest X-rays with CheXNet and feature selection algorithms, and the accuracy of the classifiers such as SVM and XGBoost reached almost 99.6% which demonstrated the power of the prediction based on the image.

Other researches were concerned with patient records and laboratory results. Models used by Alotaibi et al. [3] to identify high-risk patients early included Random Forest, SVM and Neural Networks, whereas Xiong et al. [4] concluded that Random Forest was the best in predicting the severity based on the CT scans, D-dimer levels, and neutrophil-to-lymphocyte ratios. Aljameel et al. [5] also pointed out that Remote Forest with pre-processing methods produced an accuracy of 95% and AUC of 0.99, indicating that it is a reliable method of use in real-life scenarios.

Hybrid structures based on imaging have also been researched extensively. Aswathy et al. [6] compared CT characteristics with previously trained models such as DenseNet and ResNet, identifying cases as low-, moderate-, and high-risk with an accuracy of more than 90%. Chieregato et al. [7] suggested a hybrid ML/DL model, which incorporated CT features, clinical data, and lab results, with the CatBoost AUC of the test set being 0.949. These were the methods that were predictive and interpretable to support clinical decisions.

A number of reviews have also summarized the results in this direction. Bhosale and Patnaik [8] were able to provide a systematic analysis of DL methods, mentioning the prevalence of CNN but also highlighting such issues as a lack of generalization and clinical acceptance. Alakus and Turkoglu [9] made a comparison between several deep learning methods and realized that DL models can always outperform the traditional ML in the cases where huge datasets are present. On the same note, John-Otumu et al. [10] noted the increased position of DL in healthcare prediction, which should be coupled with real-time monitoring to reinforce pandemic management systems.

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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141166

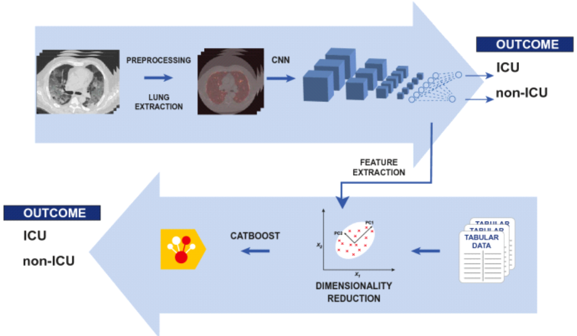

Fig. A hybrid machine learning/deep learning COVID-19 severity predictive model from CT images

The combination of these works reveals that the use of ML/DL with imaging and clinical characteristics results in the reliable and accurate prediction of COVID-19 severity. Nonetheless, the majority of the current models are based on either imaging or patient records, whereas a gap exists in the area of integrated frameworks that provide accuracy, interpretability, and clinical applicability.

II. LITERATURE SURVEY

Keerthi and Naga Raju (2025) study the subject of predicting the severity of COVID-19 cases with the help of machine learning and deep learning. The article is dedicated to examining patient data and defining the factors, which have an impact on the disease outcome. The authors compare methods using traditional machine learning models with advanced deep learning architectures to highlight both the strong and weak sides of each of the methods, when it comes to medical prediction tasks. Their paper focuses on the fact that, due to their capability to process highly dimensional, and complex data, deep learning models tend to be more accurate than the classical ones. Meanwhile, they observe the significance of explainability and reliability in practice. The study adds to the expanding body of research on healthcare analytics, with predictive models helping prevent the early identification of patients, their follow-ups, and resources when pandemics occur. [1]

Researchers are working on AI methods in order to assist the doctors to forecast severity of COVID-19 cases. This paper used the images of chest X-rays in order to create a prediction model. The CheXNet deep model and handcrafted techniques were used to extract features and the PCA and RFE methods were used to pick important features. Classifiers like XGBoost and SVM which are machine learning yielded good results with XGBoost achieving 97% accuracy. Extra Trees and SVM had near-perfect accuracy (99.6 percent) when CheXNet features were applied with RFE. The paper demonstrates that early prediction of COVID-19 severity can be offered with reliable tools through the combination of deep learning and the feature selection. [2]

This is significant because the severity of COVID-19 can be predicted at an early age to save lives, particularly in the countries that have a small medical force. In this study machine learning techniques (Artificial Neural Networks, Support Vector Machines, and Random Forest regression) were employed with the use of patient history and lab test outcomes. The findings demonstrated that these models have the potential to promptly and effectively identify patients at increased risk and enable doctors to prioritize the treatment and allocate hospital resources more effectively. [3]

Researchers examined COVID-19 patients to determine the severity of the disease upon admission to the hospital with the help of machine learning. Patient data on 287 patients, such as clinical data, lab data, and CT scans, were reviewed. Random Forest (RF), Support Vector Machine (SVM) and Logistic Regression (LR) three models were compared. Random Forest model was found to be the most accurate and sensitive. Chest CT scans, neutrophil-to-lymphocyte ratio, lactate dehydrogenase, and D-dimer levels were important features that were associated with severity. According to the study, RF can assist physicians in diagnosing at-risk patients and managing resources more effectively in a short time. [4]

The COVID-19 pandemic has brought major economic and health challenges in the world as most patients continue to die. The objective of the study was to forecast survival of COVID-19 patients based on clinical features gathered during the quarantine. Three machine learning models Logistic Regression, random forest, and XGBoost were used to analyze data of 287 patients in King Fahad University Hospital, Saudi Arabia. The data quality and imbalance were dealt with using preprocessing techniques, cross-validation, and SMOTE. Random Forest had the highest performance of 95%

DOI: 10.17148/IJARCCE.2025.141166

accuracy and 0.99 AUC. The model will be able to assist the healthcare personnel in identifying the high-risk patients at an early stage and enhance decision-making on the treatment process. [5]

The COVID-19 being a global health emergency is primarily lung-attacking, and it is hard to differentiate between the disease and other lung diseases. This paper presented the two-step approach where the first step detects the COVID-19 on the lung CT image and the second step is the classification of the severity of the illness. Pre-trained models such as AlexNet, DenseNet-201 and ResNet-50 were used to extract features and then an Artificial Neural Network was used to detect the features. The severity was determined by using the combination of CT image features and clinical data and categorizing cases as high, moderate and low risk using a Cubic SVM. The strategy reached 92% and 90% accuracy in the COVID-19 and severity classification respectively. This assists doctors to give high risk patients priority so that they can receive timely care. [6]

COVID-19 has numerous symptoms, including mild and severe respiratory failure and death. This paper created a ML/DL hybrid model to identify patients who would either need ICU services or those patients who were at risk of high mortality. They used data of 558 patients in an Italian hospital with the combination of CT scan characteristics, laboratory results, and clinical data. The extraction of features using CT was done using a 3D CNN, and the feature selection was done using Boruta and SHAP values. The test set CatBoost model gave a test set AUC of 0.949. This methodology gives physicians probabilities of outcomes and feature importance that can be interpreted and used to make clinical judgments [7].

III. PROPOSED METHOD

3.1 Data Collection: The initial one will be to collect COVID-19 related data including chest X-ray or CT scans, medical history, and laboratory test outcomes. Publicly available datasets (such as Kaggle COVID-19 Radiography Database) will be used, and the relevant features will be chosen to train and test them.
3.2 Data Pre-processing: Raw data is cleansed and ready before the application of any model. This will involve noise removal, image size normalization, missing values, dataset balancing, and data split into training set and testing set. The techniques of feature extraction are used to bring out significant information which assists in severity prediction.
3.3 Feature Optimization: Hybrid Grey Wolf Optimization and Particle Swarm Optimization are optimization techniques that are used to optimize and narrow down to the most pertinent features. The step eliminates undesired data and enhances model accuracy and performance.
3.4 Model Development: UNET (image segmentation) and ResNet50/CNN (classification) deep learning structures are used. The UNET model is used to mark the affected areas in medical images, and CNN-based models are used to assign the scan to various categories, e.g. normal, pneumonia, and COVID-19.
3.5 Hybrid Model Extension: Random Forest classifier combined with optimized features of ResNet50 is used to enhance prediction. This hybrid model takes advantage of the fact that ResNet50 is effective in features extraction and the decision-making capacity of the Random Forest, which leads to the increased accuracy.
3.6 Model Training and Testing: The model is trained on the prepared dataset, and another subset is retained in order to be tested. The models are built in Python and trained with the unknown data to test the performance. Cross-validation leads to reliability of results.
3.7 Evaluation Metrics: Accuracy, precision, recall, F1-score, and AUC are used to determine the performance of models. The visualization of the performance of the proposed approach against existing methods is done with the help of graphs and confusion matrices.
3.8 Demonstration on User Interface.
Graphical user interface (GUI) is created to render the system interactive and user friendly. The GUI will provide users with the ability to upload medical images, see distinguished areas, and get automatic disease classification and prediction outcomes.

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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141166

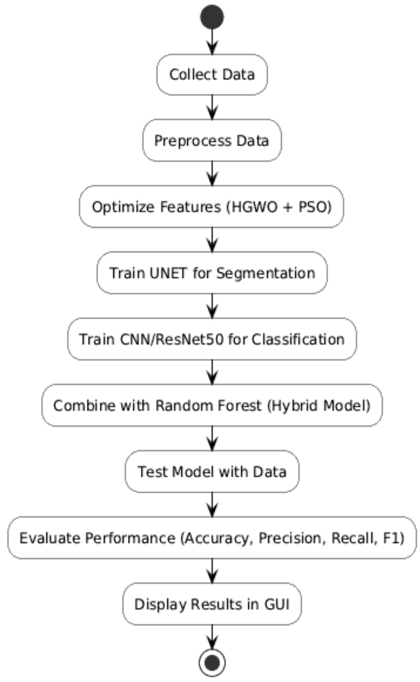

Fig.3.1 BLOCK DIAGRAM

IV. RESULTS

As extension we have extracted optimized features from train ResNet50 model and then retrained with Random Forest classifier to build a hybrid classification model. ResNet50 known for optimized features extraction and this optimized features can help other classifiers in gaining better accuracy. Random Forest trained on Hybrid ResNet50 features giving an accuracy of 100%.

We have coded this project using JUPYTER notebook and below are the code and output screens with blue colour comments

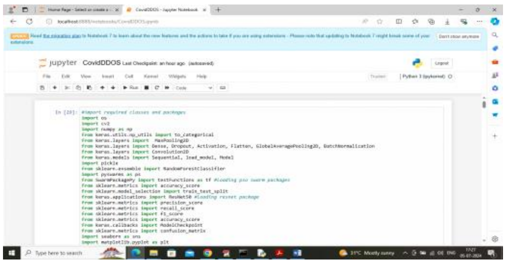

Fig. 4.1 Importing required Python classes and packages

In above screen importing required python classes and packages

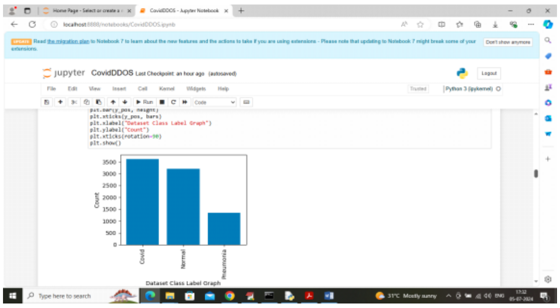

Fig. 4.2 Graph showing number of images per class label

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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141166

In above screen displaying graph of different class labels where x-axis represents label names and y-axis represents number of images under that class label

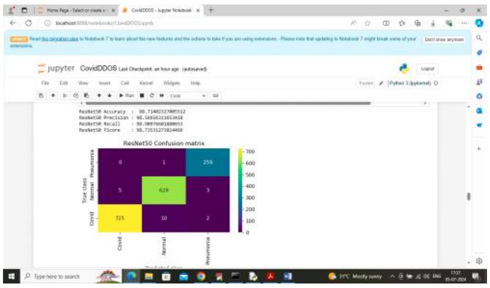

Fig. 4.3 ResNet50 model accuracy, precision, recall, F1-score with confusion matrix visualization

In above screen ResNet50 got 98.71% accuracy closer to 99% and can see other metrics like precision, recall and FSCORE. In confusion matrix graph x-axis represents ‘Predicted Labels’ and y-axis represents True Labels and then all different colour boxes in diagnol represents correct prediction count and remaining blue boxes represents incorrect prediction count which are very few

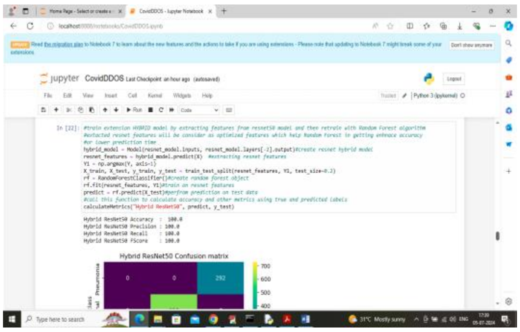

Fig. 4.4 Creation of Hybrid ResNet50 model with optimized features

In above screen creating Hybrid ResNet50 model as extension and after training on Resnet50 optimized features hybrid model got 100% accuracy

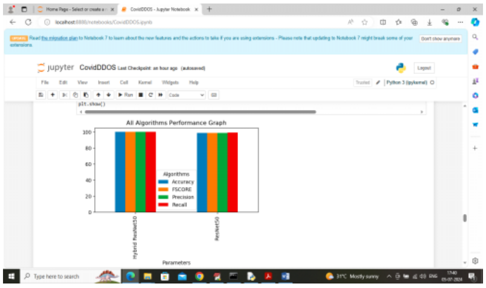

Fig. 4.5 Accuracy comparison graph of different algorithms

In above graph x-axis represents algorithm names and y-axis represents accuracy and other metrics and extension algorithm got slight high accuracy

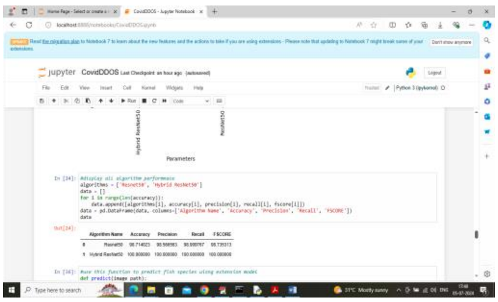

Fig. 4.6 Tabular comparison of algorithm performance metrics

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

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141166

In above screen can see both algorithm performance in tabular format

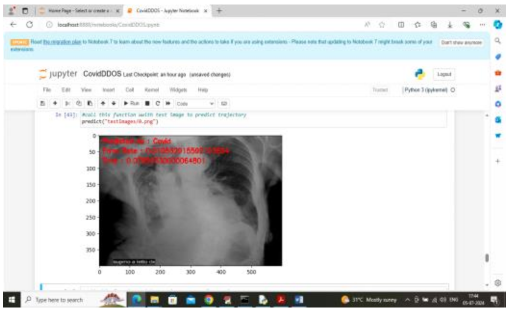

Fig. 4.7 Prediction output showing test image classified as COVID-19

In above screen calling predict function with test image path and then in RED colour text can see image predicted as ‘Covid19’ along with error rate and prediction time

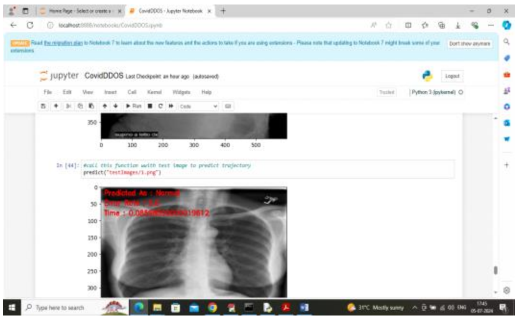

Fig. 4.8 Prediction output showing test image classified as Normal

Above image detected as Normal

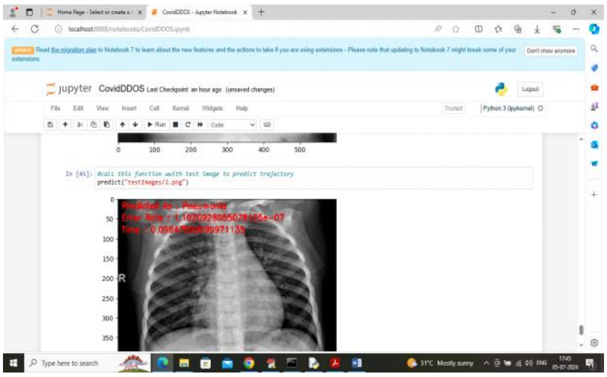

Fig. 4.9 Prediction output showing test image classified as Pneumonia

Above image detected as ‘Pneumonia’.
Similarly you can input any test image and detect and classify disease and by applying OFBC we can save image from DDOS attacks.

V. CONCLUSION

The formulated methodology shows that the combination of UNET to perform segmentation, CNN/ResNet50 to classify, and Random Forest to make hybrid decisions is an effective system in terms of predicting COVID-19 severity. The hybrid approach is more accurate and reliable, and it has less manual work as compared to the traditional models. The system is practical and can be used in practice in the real world due to the inclusion of a GUI, which can help medical professionals and patients. The model also emphasizes the role of artificial intelligence as a means of improving healthcare by providing fast, automatized, and precise diagnostic help, particularly in pandemic cases, when timely decision-making can save lives.

DOI: 10.17148/IJARCCE.2025.141166

REFERENCES

[1]. Keerthi, Naraharisetti, and A. Naga Raju. "Prediction Of COVID-19 Severity by Applying Machine and Deep Learning Techniques." International Journal of Management Research and Reviews 15, no. 2s (2025): 387-392.

[2]. Sayed, Safynaz Abdel-Fattah, Abeer Mohamed Elkorany, and Sabah Sayed Mohammad. "Applying different machine learning techniques for prediction of COVID-19 severity." Ieee Access 9 (2021): 135697-135707.

[3]. Alotaibi, Aziz, Mohammad Shiblee, and Adel Alshahrani. "Prediction of severity of COVID-19-infected patients using machine learning techniques." Computers 10, no. 3 (2021): 31.

[4]. Xiong, Yibai, Yan Ma, Lianguo Ruan, Dan Li, Cheng Lu, Luqi Huang, and National Traditional Chinese Medicine Medical Team. "Comparing different machine learning techniques for predicting COVID-19 severity." Infectious diseases of poverty 11, no. 1 (2022): 19.

[5]. Aljameel, Sumayh S., Irfan Ullah Khan, Nida Aslam, Malak Aljabri, and Eman S. Alsulmi. "Machine Learning-Based Model to Predict the Disease Severity and Outcome in COVID-19 Patients." Scientific programming 2021, no. 1 (2021): 5587188.

[6]. Aswathy, A. L., Hareendran S. Anand, and SS Vinod Chandra. "COVID-19 severity detection using machine learning techniques from CT-images." Evolutionary Intelligence 16, no. 4 (2023): 1423-1431

[7]. Chieregato, Matteo, Fabio Frangiamore, Mauro Morassi, Claudia Baresi, Stefania Nici, Chiara Bassetti, Claudio Bnà, and Marco Galelli. "A hybrid machine learning/deep learning COVID-19 severity predictive model from CT images and clinical data." Scientific reports 12, no. 1 (2022): 4329.

[8]. Bhosale, Yogesh H., and K. Sridhar Patnaik. "Application of deep learning techniques in diagnosis of covid-19 (coronavirus): a systematic review." Neural processing letters 55, no. 3 (2023): 3551-3603.

[9]. Alakus, Talha Burak, and Ibrahim Turkoglu. "Comparison of deep learning approaches to predict COVID-19 infection." Chaos, Solitons & Fractals 140 (2020): 110120.

[10]. John-Otumu, Adetokunbo MacGregor, Charles Ikerionwu, Oluwaseun Oladeji Olaniyi, Oyewole Dokun, Udoka Felista Eze, and Obi Chukwuemeka Nwokonkwo. "Advancing COVID-19 prediction with deep learning models: A review." In 2024 International conference on science, engineering and business for driving sustainable development goals (SEB4SDG), pp. 1-5. IEEE, 2024.

