Integrated Hybrid Multi-Model Ensemble Architecture for Optimized Diagnostic Precision in COVID-19 Detection via Chest X-Ray Imaging

Dr. S. Rajalakshmi[1]; Srinivasan K.[2]; Lavanya J.[2]; Pooja T. S. R.[3]; Raajalakshimi R.[3]; Sabeeha Farheen[3]; Sakthisri A.[3]; Shrmila A.[3]

[1]Associate Professor, Department of Computer Science, Sri Venkateswara College of Engineering (SVCE)
[2]Assistant Professor, Department of Computer Science, Sri Venkateswara College of Engineering (SVCE)
[3]Final Year Computer Science Students, Sri Venkateswara College of Engineering (SVCE), Sriperumbudur, Chennai.

Publication Date: 2026/03/03

Abstract: Ensemble Approach for Enhanced COVID-19 Detection using Chest X-Rays introduces an automated classification system that detects COVID-19 in chest X-rays by employing hierarchical classification and augmented images. Using a majority vote-based ensemble of five distinct supervised algorithms, this system aggregates predictions to improve decision accuracy. By recognizing unique radiographic texture patterns associated with COVID-19 through advanced feature extraction techniques like statistical texture descriptors, the model aids in early detection, addressing the urgent need for effective screening. The ensemble method enhances diagnostic accuracy and reliability by combining the strengths of multiple classifiers, each adding a unique perspective that strengthens the final prediction. Data augmentation adds robustness, compensating for variations in X-ray images and improving adaptability across diverse datasets. Additionally, identifying distinct texture patterns and refining feature extraction techniques contribute to a precise and consistent diagnostic model. This project underscores the transformative potential of machine learning-driven medical imaging, highlighting benefits such as speed, precision, and reliability. Ov erall, the proposed model not only meets the demand for efficient COVID-19 screening but also marks a significant advancement in automated infectious disease detection through medical imaging analysis.

Keywords: *COVID-19 Detection, Chest X-Ray, Ensemble Approach, Supervised Algorithms, Feature Extraction, Data Augmentation, Medical Imaging, Diagnostic Accuracy.*

How to Cite: Dr. S. Rajalakshmi; Srinivasan K.; Lavanya J.; Pooja T. S. R.; Raajalakshimi R.; Sabeeha Farheen; Sakthisri A.; Shrmila A. (2026) Integrated Hybrid Multi-Model Ensemble Architecture for Optimized Diagnostic Precision in COVID-19 Detection via Chest X-Ray Imaging. *International Journal of Innovative Science and Research Technology,* 11(2), 2383-2400. https://doi.org/10.38124/ijisrt/26feb917

I. INTRODUCTION

➢ *Machine Learning*

Machine learning (ML) is a branch of artificial intelligence (AI) that focuses on enabling computers to learn from data and make decisions or predictions without explicit programming for each specific task. At its core, machine learning revolves around algorithms that identify patterns, relationships, and structures within data, allowing the computer to "learn" from past experiences or examples. This ability to learn autonomously and improve over time has made machine learning a powerful tool for solving complex problems in various domains, including healthcare, finance, manufacturing, and entertainment.

The development of machine learning is closely tied to advancements in computing power, data storage, and algorithm design, which have made it possible to process and analyze massive amounts of data at high speeds. ML methods are categorized into several types based on the learning approach they use, including supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning.

The success of machine learning relies on data quality and the choice of algorithms, which vary depending on the problem at hand. Some popular algorithms include decision trees, support vector machines, and neural networks. Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), excel in handling unstructured data like images, video, and text,

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

making them integral to applications like image recognition, natural language processing, and voice synthesis. With these advanced techniques, machine learning models can achieve remarkable accuracy and robustness, even in highly dynamic and complex environments.

Machine learning models must undergo a training process in which the model learns to recognize patterns in the training data and generalize them to new, unseen data. This process often involves iterative optimization of model parameters to minimize prediction errors and improve accuracy. One of the challenges in machine learning is finding the right balance between a model that learns well from the training data and one that can generalize to real-world data, avoiding issues such as overfitting or underfitting. Overfitting occurs when a model learns too much from the training data, capturing noise and making it less effective on new data, while underfitting happens when the model fails to capture the underlying patterns.

Machine learning is continuously evolving, with researchers and practitioners pushing the boundaries of what is possible. Recent developments include transfer learning, where a model trained on one task is adapted for a related task, and federated learning, which enables collaborative model training across decentralized devices without sharing raw data, preserving privacy. The rise of big data and improvements in hardware, especially Graphics Processing Units (GPUs) and specialized chips, have further accelerated machine learning applications, enabling faster and more efficient model training.

Machine learning has transformed many industries, offering innovative solutions for various tasks. In healthcare, ML models are applied to medical imaging analysis, disease prediction, and drug discovery. For example, algorithms can analyze chest X-rays to detect conditions such as pneumonia or COVID-19, helping radiologists diagnose diseases faster and more accurately. In finance, ML is used for fraud detection, risk assessment, and algorithmic trading.

➢ *Classifications*

Predictions for COVID-19, pneumonia, and normal chest X-rays are derived through a sophisticated ensemble approach that aggregates the results of multiple classification models. Each image is analyzed using six supervised algorithms—Naive Bayes, Decision Tree, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Convolutional Neural Network (CNN)—as well as pre-trained deep learning architectures, and VGG16 which are fine-tuned to detect radiographic anomalies specific to these conditions. Through majority voting, the system determines the final classification label, leveraging the strengths of each model to increase diagnostic accuracy and reduce misclassifications.

- *COVID-19*

Ensemble Approach for Enhanced COVID-19 Detection using Chest X-Rays introduces an automated classification system that detects COVID-19 in chest X-rays by employing hierarchical classification and augmented images. Using a majority vote-based ensemble of five distinct supervised algorithms, this system aggregates predictions to improve decision accuracy. By recognizing unique radiographic texture patterns associated with COVID-19 through advanced feature extraction techniques like statistical texture descriptors, the model aids in early detection, addressing the urgent need for effective screening.

The ensemble method enhances diagnostic accuracy and reliability by combining the strengths of multiple classifiers, each adding a unique perspective that strengthens the final prediction. Data augmentation adds robustness, compensating for variations in X-ray images and improving adaptability across diverse datasets. Additionally, identifying distinct texture patterns and refining feature extraction techniques contribute to a precise and consistent diagnostic model. This project underscores the transformative potential of machine learning-driven medical imaging, highlighting benefits such as speed, precision, and reliability. Overall, the proposed model not only meets the demand for efficient COVID-19 screening but also marks a significant advancement in automated infectious disease detection through medical imaging analysis.

- *Pneumonia*

Pneumonia cases, which exhibit distinct radiographic signs, are similarly identified with high accuracy by the ensemble. Pneumonia images typically show consolidation, which the model can distinguish from COVID-19-associated patterns through variations in texture and opacity distribution. Models like the Decision Tree and KNN, which are effective at categorizing more structured image data, provide insights into these features, while the CNN contributes by learning localized features across various augmented samples. The integration of multiple models helps the system maintain a balanced prediction capability for pneumonia cases, minimizing both false positives and false negatives, which is essential for accurate diagnostics in the healthcare domain.

- *Normal*

For chest X-rays classified as normal, the ensemble model evaluates the absence of abnormalities, such as irregular opacities or consolidations. Normal images lack the distinct texture disruptions present in COVID-19 and pneumonia cases, and the ensemble's majority voting ensures that these cases are not misclassified. The SVM and Naive Bayes models add value by confirming the normal class through statistical patterns and overall image consistency. The CNN and transfer learning models, by processing large volumes of healthy lung X-rays, further enhance the model's ability to distinguish between normal and pathological findings accurately. The ensemble's robust training on augmented data enables it to identify normal cases across diverse patient samples, ensuring reliable differentiation even when facing variations in image quality or demographic differences.

➢ *Ensemble Learning*

Ensemble learning is to enhance the accuracy and reliability of COVID-19 detection from chest X-rays by leveraging the strengths of multiple machine learning and

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

deep learning models. This approach combines predictions from diverse classifiers, each trained to recognize different radiographic patterns, into a unified decision through majority voting. By mitigating individual model biases and reducing misclassifications, ensemble learning increases diagnostic consistency and improves the identification of COVID-19, pneumonia, and normal cases. Integrating traditional algorithms with deep learning architectures, along with data augmentation to simulate variability, creates a robust framework that adapts to diverse datasets. Ultimately, ensemble learning significantly strengthens the model, enabling a reliable and automated tool for infectious disease detection.

- *Convolutional Neural Networks (CNNs)*

Convolutional Neural Networks (CNNs) are a specialized type of deep learning model designed primarily for processing structured grid data, such as images. They have emerged as the leading architecture for a variety of tasks in computer vision, including image classification, object detection, and image segmentation.

CNNs are inspired by the biological processes underlying human vision, where the brain recognizes patterns through a series of hierarchical feature extractions. By automatically learning spatial hierarchies of features from input images, CNNs significantly reduce the need for manual feature engineering and can efficiently identify complex patterns, making them particularly effective for analyzing visual data.

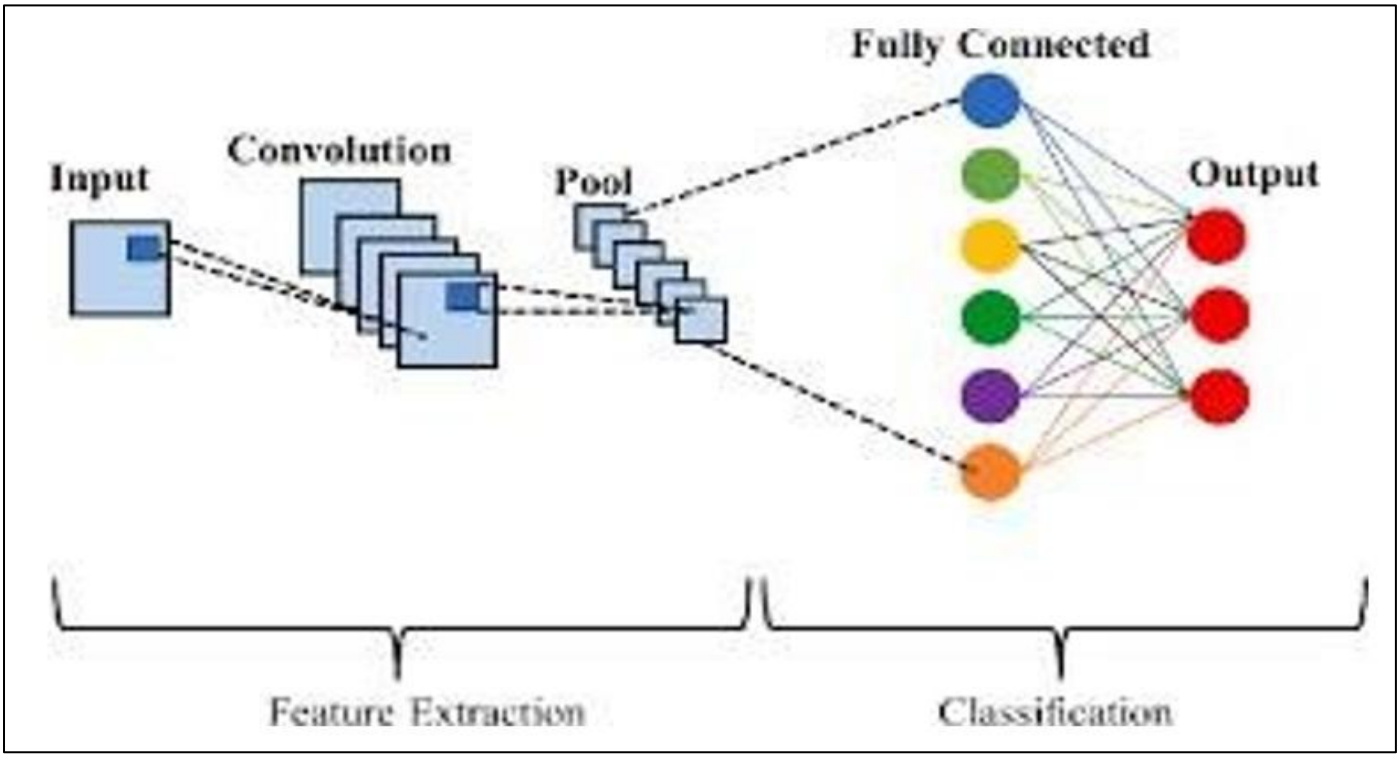

Fig 1 Basic Architecture of Convolutional Neural Network

The strength of CNNs lies in their ability to capture the spatial relationships between pixels in an image. Unlike traditional fully connected neural networks that treat input data as a flat vector, CNNs maintain the two-dimensional structure of images. This structural awareness enables CNNs to recognize patterns regardless of their position in the image.

The architecture of a CNN is specifically tailored to leverage this property, consisting of various layers that work collaboratively to extract features and make predictions. As CNNs have evolved, they have incorporated deeper architectures and more sophisticated training techniques, allowing them to achieve state-of-the-art performance across numerous applications.

The architecture of CNNs is built around the concept of convolution, a mathematical operation that applies a filter (or kernel) over the input image to produce a feature map. This feature map highlights specific patterns within the image, such as edges, corners, and textures. The primary advantage of using convolutional layers is their ability to preserve the spatial relationships between pixels while reducing the number of parameters in the model, which makes CNNs computationally efficient and less prone to overfitting.

- *Decision Tree (DT)*

Decision trees are a prominent and versatile machine learning model that has gained significant traction across various fields, including healthcare, finance, and marketing. They offer a straightforward yet powerful approach to classification and regression tasks by representing decisions and their possible consequences in a tree-like structure.

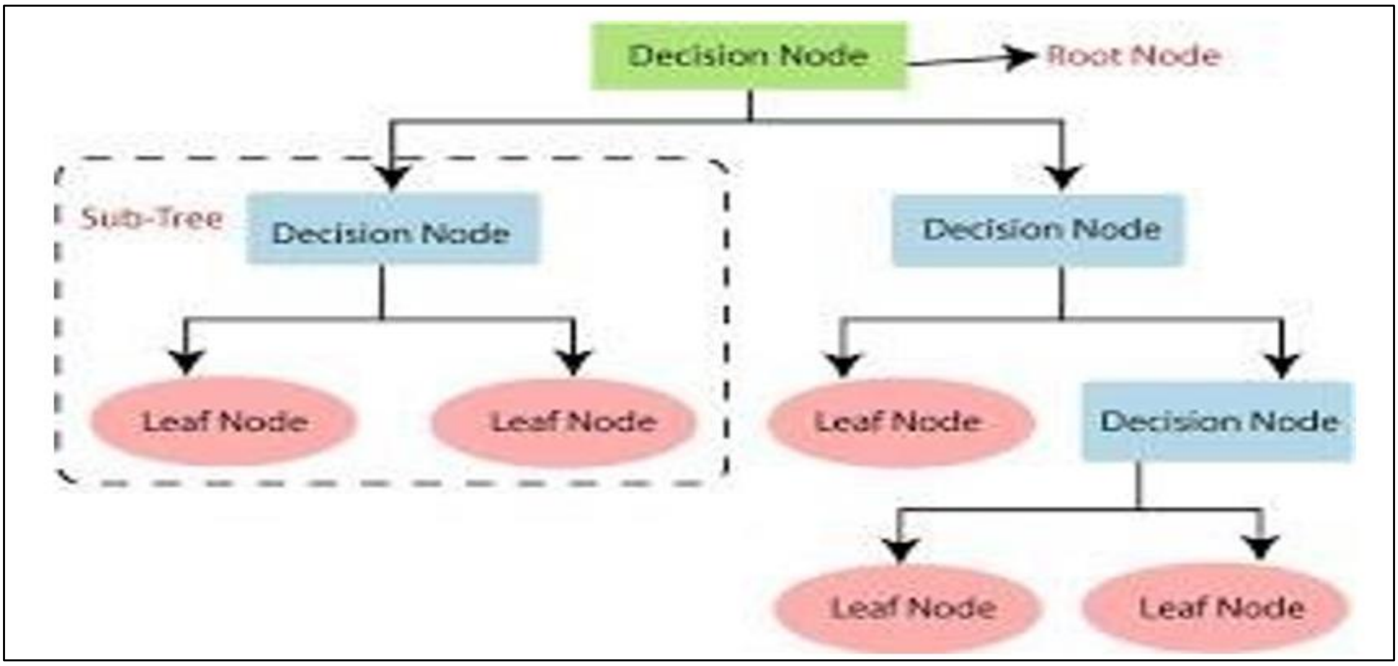

Fig 2 Block Diagram of Decision Tree

The construction of a decision tree begins with the selection of the most significant feature for splitting the dataset. This selection is guided by specific criteria that measure the quality of the split, such as Gini impurity, which quantifies how often a randomly chosen element would be incorrectly labeled if it were randomly labeled according to the distribution of labels in the subset, or information gain, which measures the reduction in entropy or uncertainty achieved by a split.

The goal is to partition the data into subsets that are as homogeneous as possible with respect to the target variable. The splitting process is recursive; the tree continues to grow by making further splits on the resulting subsets until a stopping criterion is reached. These criteria may include a maximum tree depth, a minimum number of samples required to split a node, or achieving a certain level of purity in the leaf nodes.

One of the major advantages of decision trees is their ability to handle both numerical and categorical data seamlessly. They do not require extensive data preprocessing, such as normalization or standardization. Decision trees are robust to outliers, as they partition the data based on feature values rather than being influenced by extreme values.

Decision trees are a foundational machine learning model characterized by their simplicity, interpretability, and versatility. They effectively capture both linear and non-linear relationships within the data, making them suitable for a broad spectrum of applications, including medical diagnosis, financial forecasting, and risk assessment.

- *K-Nearest Neighbor (KNN)*

The K-Nearest Neighbor (KNN) algorithm is a simple yet powerful machine learning technique used for classification and regression tasks. It is categorized as a lazy learning algorithm because it does not explicitly learn a model during the training phase; instead, it memorizes the training instances and makes predictions based on the proximity of these instances to the input data points.

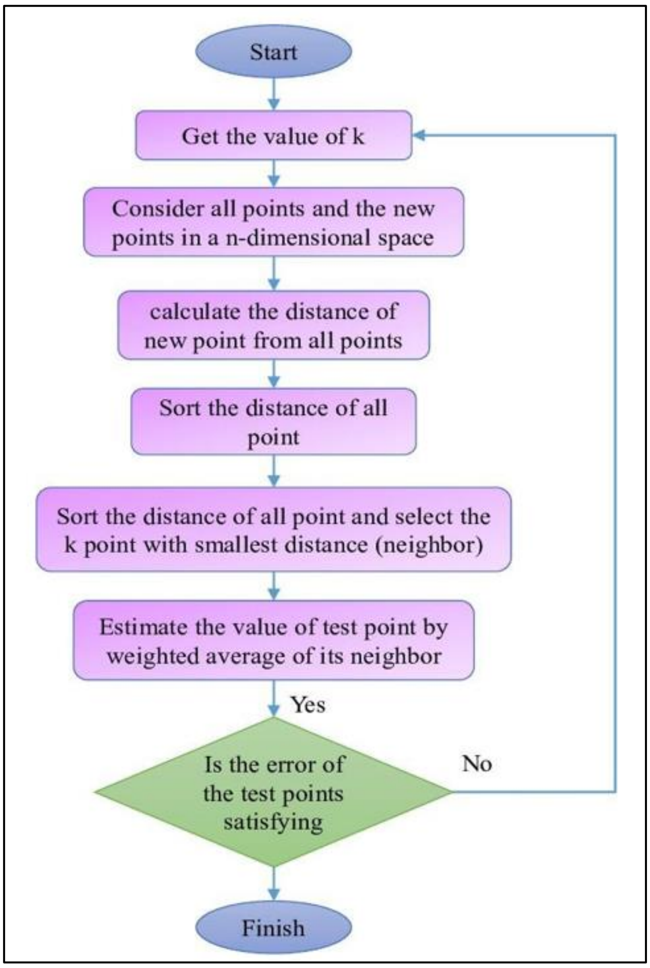

Fig 3 Architecture of K-Nearest Neighbor

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

KNN operates on the principle that similar instances tend to be close to each other in the feature space, making it particularly effective for problems where the decision boundary is irregular and not easily defined by a linear model.

The KNN algorithm works by determining the distance between a query point and all the training instances in the dataset. The most common distance metric used is the Euclidean distance, although other metrics such as Manhattan distance and Minkowski distance can also be applied depending on the context. Once the distances are calculated, the algorithm identifies the 'k' nearest neighbors, where 'k' is a user-defined parameter. The prediction for the query point is then made based on the majority class among these neighbors.

KNN's intuitive approach of leveraging proximity to make predictions makes it suitable for applications ranging from image recognition to medical diagnosis. In the context of detecting diseases such as COVID-19 and pneumonia using chest X-rays, KNN can be employed to classify images based on features derived from the X-ray data.

- *Naive Bayes (NB)*

The Naive Bayes model is a family of probabilistic classifiers based on Bayes' theorem, which applies a strong (naive) assumption of independence among features. Despite its simplicity, Naive Bayes has proven to be a highly effective model for various classification tasks, especially in scenarios with high- dimensional data and limited computational resources.

This characteristic makes it particularly suitable for applications in fields like text classification, spam detection, and medical diagnosis, where understanding the probabilistic relationships among features can lead to insightful predictions.

The underlying principle of the Naive Bayes classifier is Bayes' theorem, which describes the relationship between the conditional probabilities of events. The formula can be expressed as:

$$P(Y|X) = \frac{P(X|Y) \cdot P(Y)}{P(X)}$$

✓ *In the Naive Bayes Formula:*

- X represents the set of features or the observed data. In your case, with COVID-19 detection, XXX could represent the features extracted from the chest X-ray images, such as specific radiographic texture patterns or pixel intensities.
- Y represents the class label or the category to be predicted. Here, YYY would correspond to the classification labels, such as "COVID-19 Positive" or "COVID-19 Negative."

Thus, P(Y|X) P(Y|X) P(Y|X) is the probability of the class (COVID-19 status) given the observed features (X-ray characteristics).

In a Naive Bayes classifier, the key assumption is that each feature in the dataset contributes independently to the probability of the outcome. This allows for a simplified calculation of the likelihood P(X|Y) as the product of the probabilities of the individual features given the class:

$$P(X|Y) = P(x_1|Y) \cdot P(x_2|Y) \cdot \ldots \cdot P(x_n|Y)$$

Where x1, x2…, Xn are the features of the input data. This simplification leads to a very efficient algorithm, as it significantly reduces the complexity of the calculations involved, making Naive Bayes particularly scalable for large datasets.

- *Support Vector Machine (SVM)*

Support Vector Machines (SVM) are a class of supervised learning algorithms that are widely used for classification and regression tasks. Developed by Vladimir Vapnik and his colleagues in the 1990s, SVMs have gained prominence due to their effectiveness in handling high-dimensional spaces and their robust performance in a variety of applications.

The SVM algorithm operates by mapping the input data into a high-dimensional space using a kernel function. This transformation allows SVM to tackle non-linear decision boundaries effectively. The main objective of the algorithm is to maximize the margin, which is the distance between the hyperplane and the nearest data points from either class.

These nearest points are called support vectors, and they play a crucial role in defining the optimal hyperplane. The optimization problem can be formulated mathematically as:

$$\text{maximize} \quad \frac{2}{||w||} \quad \text{subject to} \quad y_i(w \cdot x_i + b) \geq 1 \quad \forall i$$

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

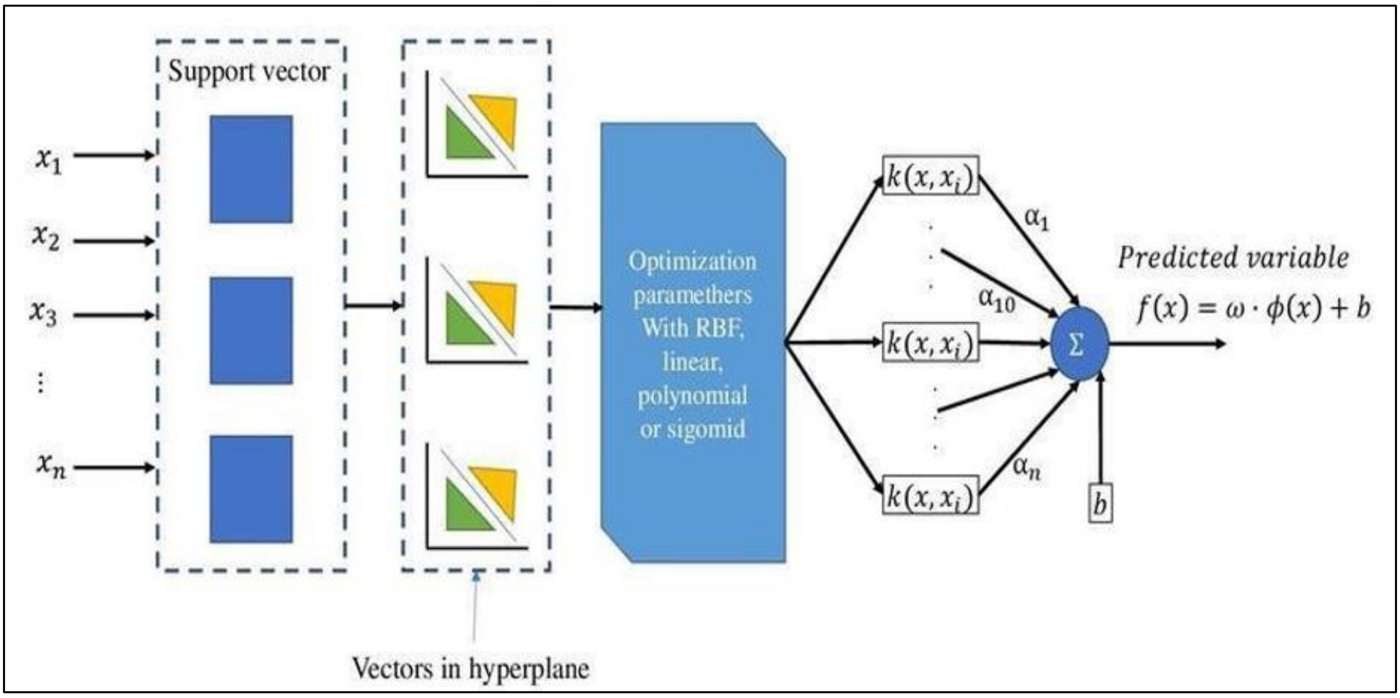

Fig 4 Architecture of Support Vector Machine

SVMs can utilize different kernel functions to adapt to the complexity of the dataset. The choice of kernel allows SVM to create non-linear decision boundaries, enabling it to separate classes that are not linearly separable in the original feature space.

One of the significant advantages of SVM is its effectiveness in high-dimensional spaces, which makes it suitable for applications with many features, such as image data and gene expression data. Furthermore, SVM is robust to overfitting, especially when the number of dimensions exceeds the number of samples.

- *Visual Geometry Group (VGG16)*

VGG16, introduced by the Visual Geometry Group is a convolutional neural network (CNN) architecture that has gained widespread recognition in the field of deep learning for its performance in image classification tasks. The architecture is characterized by its simplicity and depth, consisting of 16 weight layers, which include 13 convolutional layers and three fully connected layers.

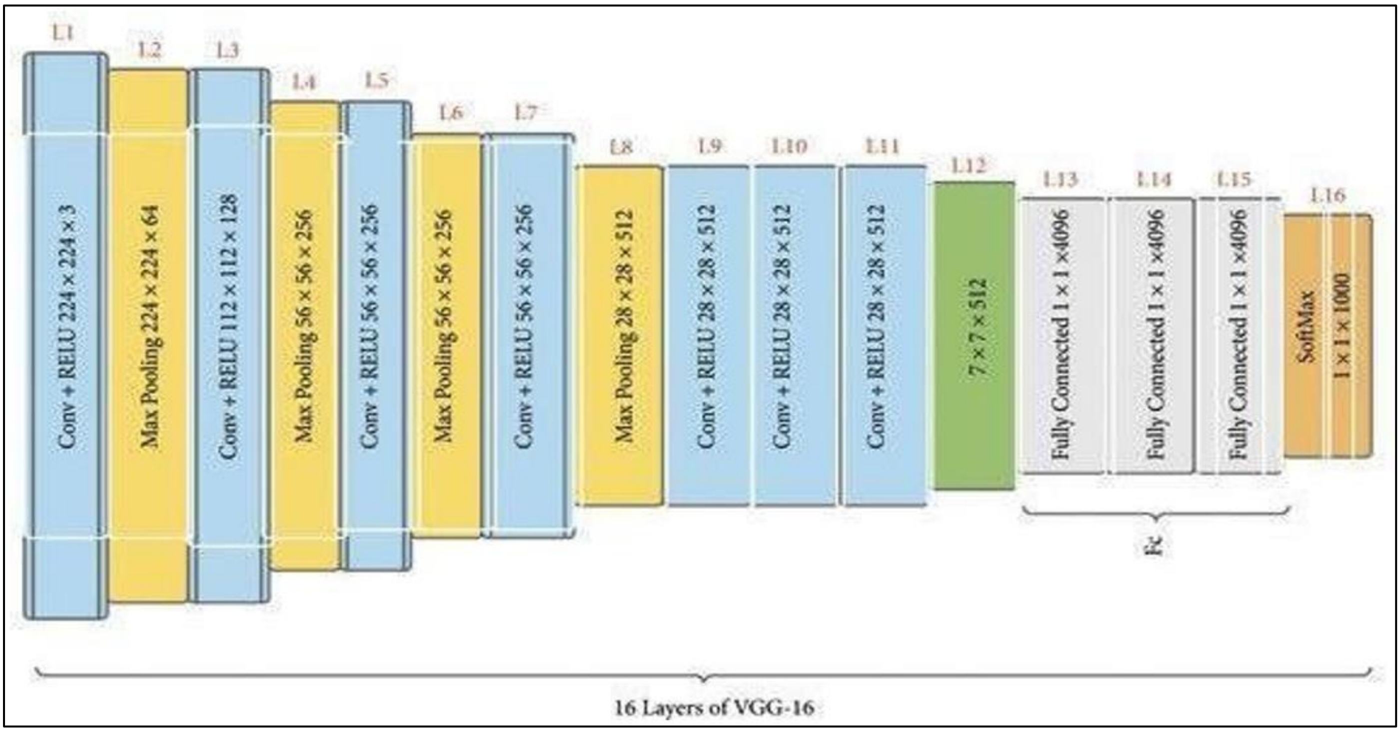

Fig 5 Basic Architecture of VGG16 Model

VGG16, introduced by the Visual Geometry Group at the University of Oxford in their 2014 paper, is a convolutional neural network (CNN) architecture that has gained widespread recognition in the field of deep learning for its performance in image classification tasks. The architecture is characterized by its simplicity and depth, consisting of 16 weight layers, which include 13 convolutional layers and three fully connected layers.

The VGG16 model is designed with small 3x3 convolutional filters stacked on top of each other, enabling the network to learn complex patterns while maintaining a manageable number of parameters. The model also employs max-pooling layers to progressively reduce the spatial dimensions of the input volume, allowing for higher-level feature extraction as the data passes through the network.

VGG16 has achieved remarkable results on the ImageNet dataset, significantly contributing to the advancement of image classification and object detection tasks. VGG16's performance has led to its adoption in numerous applications beyond image classification, including medical image analysis, where it has been effectively utilized to identify diseases such as pneumonia and COVID-19 from chest X-ray images.

- *Problem Statement*

Early and accurate detection of COVID-19 is crucial for effective disease management, as it enables timely interventions and reduces the spread of the virus. Chest X-rays present a promising and accessible modality for rapid screening due to their widespread availability and low cost compared to other imaging techniques. However, individual deep learning models for COVID-19 detection often face limitations, primarily stemming from data variability, model complexity, and the inherent challenges of medical imaging. To address these issues, this project aims to develop a robust ensemble learning framework that enhances the accuracy of COVID-19 detection using chest X-ray images.

By integrating the strengths of multiple deep learning models, the ensemble approach seeks to mitigate the weaknesses of individual models and improve overall diagnostic performance. This method not only enhances detection accuracy but also aims to provide a more reliable tool for healthcare professionals in COVID-19 triage and patient care. Ultimately, the project's goal is to contribute to more effective management of COVID-19 through improved screening processes, enabling quicker responses and better allocation of healthcare resources during the ongoing pandemic.

II. LITERATURE REVIEW

Aleka Melese Ayalew, Ayodeji Olalekan Salau, Yibeltal Tamyalew, Bekalu Tadele Abeje, and Nigus Woreta (2023) explores the use of deep learning for detecting COVID-19 in chest X-ray images. Recognizing the need for rapid and reliable diagnostic tools amid the pandemic, the researchers propose a system that leverages advanced machine learning models to automatically identify radiographic features associated with COVID-19. Their approach involves analysing texture and visual patterns specific to the disease, which helps in distinguishing COVID-19 from other lung conditions. The model is trained on a large dataset of X-ray images to improve its diagnostic accuracy and reliability. This automated system aims to assist healthcare professionals in early detection, offering an efficient and accessible solution for hospitals and clinics, especially in resource-limited settings. The study highlights the potential of AI-driven tools in enhancing medical imaging diagnostics.

Sreeparna Das, Ishan Ayus, and Deepak Gupta (2023) examine various computational methods used to identify COVID-19. The review discusses how machine learning (ML) and deep learning (DL) algorithms have been employed to enhance the diagnostic process through image-based and non-image-based approaches. The authors analyze key ML and DL models, such as convolutional neural networks (CNNs), which are highly effective in processing chest X-rays and CT scans to detect COVID-19 indicators. They also highlight challenges, such as the need for large datasets and model interpretability, that affect the deployment of AI-based tools in real-world medical settings. The review underscores the potential of ML and DL in improving COVID-19 diagnosis accuracy and efficiency, promoting advancements in healthcare technology for pandemic response and future applications.

Lin Zou, Han Leong Goh, Charlene Jin Yee Liew, Jessica Lishan Quah, and colleagues (2023) propose an Ensemble Image Explainable AI (XAI) algorithm designed to diagnose severe community-acquired pneumonia (SCAP) and COVID-19-related respiratory infections. This system combines multiple AI models to analyze medical images, providing robust and interpretable predictions that can help medical professionals in early diagnosis and management of severe respiratory infections. By using an ensemble approach, the algorithm aggregates predictions from various models to improve diagnostic accuracy, offering a more comprehensive analysis of radiographic data. Additionally, the XAI component enables the model to present visual explanations of its predictions, which is particularly valuable in the medical field where interpretability is critical. This method enhances clinicians' confidence in AI-driven results, allowing them to better understand the reasoning behind predictions, ultimately supporting more informed medical decision-making.

Haval I. Hussein, Abdulhakeem O. Mohammed, Masoud M. Hassan, and Ramadhan J. Mstafa (2023) developed lightweight, deep convolutional neural network-based models for early detection of COVID-19 from chest X-ray images. This approach addresses the critical need for rapid, reliable COVID-19 screening, particularly in areas with limited healthcare resources. By using CNNs, the researchers created a model that efficiently analyzes X-ray images to identify COVID-19's unique radiographic patterns, providing a tool that balances high accuracy with low computational demand. Their lightweight model is designed to operate on systems with limited processing power, making it feasible for widespread application in clinics and hospitals.

This research offers a significant contribution to automated COVID-19 diagnostics, showing potential to assist healthcare providers in quick and effective screening, ultimately reducing the burden on healthcare systems and aiding in managing the spread of the virus.

Vandana Bhattacharjee, Ankita Priya, Shamama Anwar (2023) presents a novel deep learning model designed to detect COVID-19 from chest X-ray images. The DeepCOVNet model utilizes convolutional neural networks to analyze X-ray images for signs of COVID-19, aiming to enhance diagnostic accuracy and speed. The study highlights the importance of rapid and reliable detection methods during the pandemic, especially in resource-limited settings. The authors evaluate the model's performance against existing methods, demonstrating its effectiveness in distinguishing COVID-19 cases from other pneumonia types and healthy individuals. Through rigorous experimentation, the model shows promising results, potentially aiding healthcare professionals in timely decision-making and improving patient outcomes. Overall, this research contributes to the field of medical imaging and emphasizes the role of artificial intelligence in combating infectious diseases.

Truong Dang John McCall, Eyad Elyan, Carlos Francisco Moreno-Garcia (2024) presents a novel approach to medical image segmentation using deep learning. The authors propose a two-layer ensemble model that combines multiple deep learning architectures to improve segmentation accuracy in medical imaging applications. By leveraging the strengths of different models, the ensemble approach enhances the robustness and reliability of the segmentation results. The first layer consists of various deep learning models that independently process the images, while the second layer aggregates their outputs to produce a more precise final segmentation. This method aims to address challenges such as variability in imaging modalities and patient anatomy. The study demonstrates that the two- layer ensemble model outperforms individual models in terms of segmentation performance, making it a promising solution for applications in medical diagnostics and treatment planning.

Talib Iqball and M. Arif Wani (2023) presents a novel approach to enhance image classification performance. It introduces a weighted ensemble model that combines multiple classifiers, leveraging their strengths to improve accuracy and robustness. The authors emphasize the importance of selecting appropriate weights for each classifier based on their performance, allowing the ensemble to focus more on the most reliable models. The study showcases experiments conducted on various image datasets, demonstrating that the weighted ensemble approach outperforms individual classifiers and traditional ensemble methods. This method effectively reduces classification errors and enhances generalization across diverse image categories. By employing this technique, the authors aim to address common challenges in image classification, such as variations in image quality and complexity, thereby contributing valuable insights to the field of information technology and computer vision.

Baijnath Kaushik, Akshma Chadha, and Reya Sharma (2023) investigate various machine learning models to assess their effectiveness in predicting COVID-19 outcomes. The study focuses on evaluating different algorithms, including decision trees, support vector machines, and neural networks, to determine their accuracy and reliability in diagnosing and prognosticating the disease. By employing a comprehensive dataset of COVID-19 cases, the authors analyze the performance metrics of each model, such as precision, recall, and F1-score, to identify the most effective approaches for predicting patient outcomes. The findings highlight the strengths and weaknesses of each model, ultimately suggesting that ensemble methods may offer superior performance due to their ability to combine multiple learning strategies. This research contributes valuable insights into the use of artificial intelligence in healthcare, particularly in improving decision-making processes related to COVID-19 prognosis.

III. METHODOLOGY

- *Objective*

The objective of the ensemble approach for enhanced COVID-19 detection using chest X-rays is to improve diagnostic accuracy by combining multiple machine learning models. By leveraging diverse algorithms, this method aims to capture various features from the X-ray images, facilitating more robust and reliable predictions. The ensemble model seeks to minimize false negatives and false positives, ultimately enhancing the ability to identify COVID-19 infections accurately. Additionally, the approach aims to provide a quick and effective diagnostic tool that can assist healthcare professionals in making informed decisions, thereby improving patient outcomes and resource allocation during the pandemic.

- *Proposed Architecture*

The architecture is structured into two main pipelines, each serving distinct roles in achieving robust and accurate COVID-19 detection in chest X-rays. The first pipeline is the Training and Testing Model, which is focused on building and refining the ensemble of classifiers using labelled training data. The second pipeline is the Validation Model, dedicated to assessing the performance and reliability of the trained ensemble on new, unseen data. These pipelines interact through a classification mechanism, where the final decision is based on the majority vote among individual classifiers.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

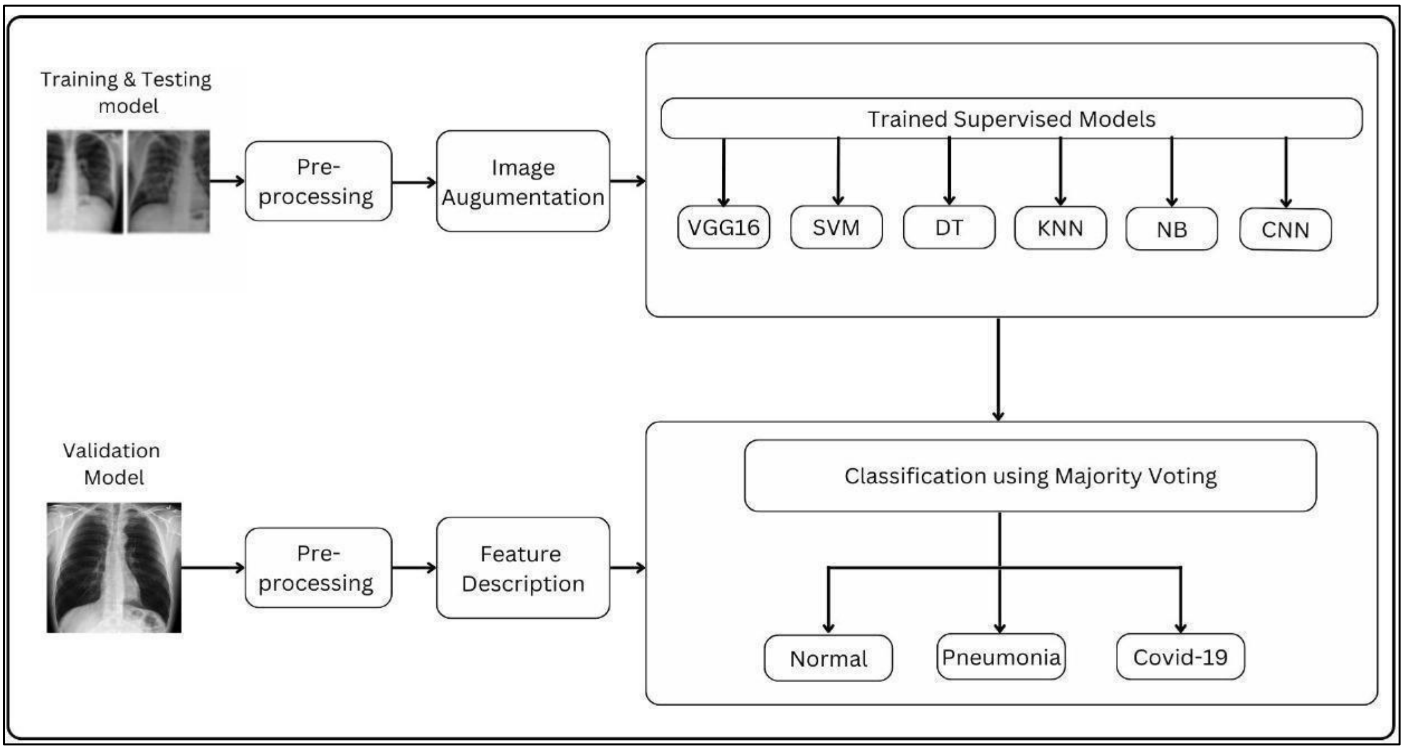

Fig 6 Proposed Architecture

The Training and Testing Model pipeline begins by taking chest X-ray images as input. These images can be from patients with different conditions such as normal, pneumonia, or COVID-19. The first step in this pipeline is Pre-processing, which is critical for preparing raw images to be more suitable for model training. During pre-processing, operations like noise reduction are applied to remove artifacts and ensure the image clarity. Images are also resized to a standard dimension, allowing for uniformity in input size across all models. Additionally, normalization is often performed to scale pixel values, which helps in stabilizing and accelerating the training process by ensuring that input features have a similar data distribution.

After pre-processing, the images proceed to the Image Augmentation step. Data augmentation is a crucial technique, especially when the dataset is limited, as it artificially increases the diversity of the training data without actually collecting new samples. Various transformations, such as rotation, horizontal and vertical flipping, zooming, and brightness adjustments, are applied to the images. These transformations introduce variations that the model might encounter in real-world scenarios, such as slight changes in orientation or lighting. By exposing the model to these variations during training, augmentation helps improve the model's generalization ability, making it more robust and effective in handling unseen images.

Following data augmentation, the augmented images are input into a set of Trained Supervised Models, creating an ensemble of classifiers. This ensemble includes both deep learning models and traditional machine learning algorithms, each bringing unique strengths to the classification task. For instance, VGG16 is a deep convolutional neural network architecture known for its effectiveness in capturing complex patterns in image data. It is a deep but straightforward CNN architecture with a uniform design, capturing fine-grained details from the images. Alongside these deep networks, machine learning models like Support Vector Machine, Decision Tree, K-Nearest Neighbour, Naive Bayes, and a simpler CNN are also trained. Each model learns independently, identifying patterns associated with normal, pneumonia, or COVID-19 cases based on its specific strengths.

The predictions from these models are then aggregated in the Classification using Majority Voting module. Majority voting is an ensemble technique where each model in the ensemble votes for a class label, and the label that receives the most votes becomes the final prediction. This approach leverages the diversity of the models to enhance the overall classification accuracy, as it reduces the risk of individual model biases or errors affecting the outcome. Majority voting is particularly useful in this context, as it allows the ensemble to produce a more reliable decision by combining the perspectives of multiple models. This method is expected to improve diagnostic accuracy by ensuring that the majority decision aligns with the most consistent patterns recognized across different models.

This final step categorizes each X-ray image into one of three possible outcomes: Normal, Pneumonia, or COVID-19. By analyzing how well the predictions align with the validation data, the architecture can assess the performance, accuracy, and reliability of the ensemble. The validation step also provides insights into potential areas for improvement, such as enhancing specific models in the ensemble or refining pre-processing and feature extraction techniques to better handle variations in image quality or other confounding factors.

This architecture, with its integration of data augmentation, diverse model ensemble, and majority voting mechanism, aims to enhance the detection of COVID-19 from chest X-rays. The inclusion of a robust validation process ensures that the trained model is not only accurate on the training data but also generalizable and reliable when applied to new cases, making it a comprehensive solution for automated COVID-19 detection.

IV. IMPLEMENTATION MODULES

➢ *Datasets*

The bar plot depicts the distribution of images across the three categories: Covid, Pneumonia, and Normal within the training set. A notable observation is the significant imbalance in the number of images between the categories.

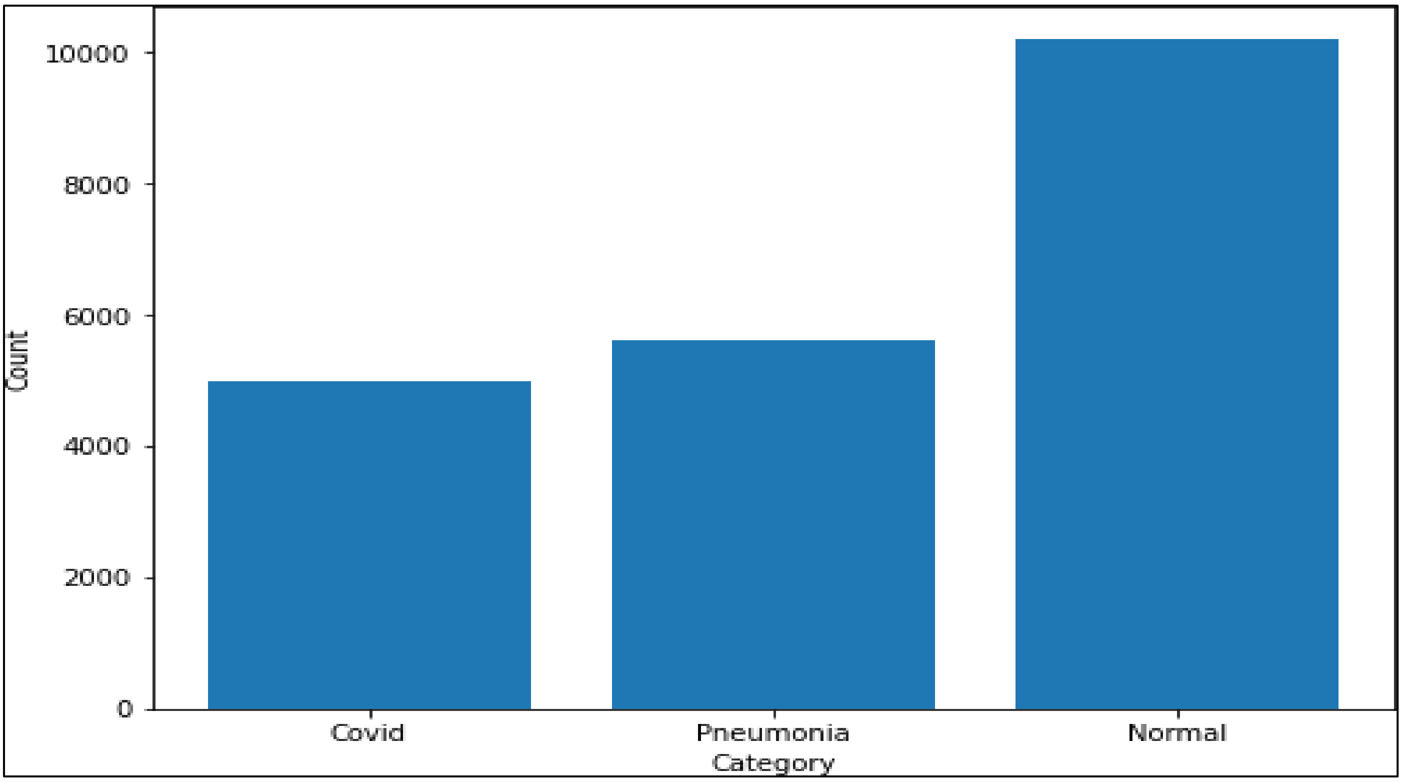

Fig 7 Number of Images Per Category in Training Set

The Normal category, with the highest count, is followed by the Pneumonia category. The Covid category, however, has the least number of images, approximately half the count of the Pneumonia category.

This class imbalance poses a potential challenge during model training, as the model might become biased towards the majority classes. To mitigate this issue, various techniques can be employed. Data augmentation, a method that artificially increases the size of the dataset by generating new, modified images, can be applied to the underrepresented classes. Additionally, class weighting can be used to assign higher importance to samples from the minority classes, thereby balancing their influence during training.

The quality of the images within each category is another crucial factor to consider. Ensuring that the images are clear, well-labelled, and diverse will contribute to a more robust model. It is essential to avoid any biases in the dataset, such as variations in image resolution, lighting conditions, or patient demographics, as these can negatively impact the model's generalizability.

By addressing the class imbalance and ensuring data quality, we can improve the model's ability to accurately classify images, even in the presence of challenging and diverse real-world scenarios.

➢ *Data Loading and Preprocessing*

The implementation begins with loading and preprocessing the chest X-ray images, critical steps in preparing the data for machine learning and deep learning models. Images are loaded from directories corresponding to the labels (Covid, Pneumonia, and Normal), resized to a uniform size of 128x128 pixels, and converted to a NumPy array format. This standardization helps ensure that all images are uniformly processed across the various models. The dataset is then split into training and testing sets, followed by normalization to scale pixel values between 0 and 1, aiding in faster convergence and improved model performance.

➢ *Label Encoding*

After loading the images, the next step is encoding the categorical labels into numerical form. This encoding step maps each label to an integer (e.g., 0 for Covid, 1 for Pneumonia, and 2 for Normal), enabling compatibility with the models. Using a dictionary to map the label names to integers simplifies the process and allows the models to interpret these categories for classification tasks effectively.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

➢ *Data Augumentation*

Data augmentation is used to artificially expand the dataset and improve model generalization by introducing variations in the images, such as rotations, shifts, shears, and flips. Using TensorFlow's `ImageDataGenerator`, these transformations create more diverse training samples, reducing the risk of overfitting. The augmented images are used in the CNN, and VGG16 training processes to help the models generalize better to unseen data, enhancing their robustness in real-world scenarios.

➢ *Training Models for Chest X-Ray Classification*

This section outlines the training methodologies used for both traditional machine learning models and deep learning architectures, including a custom CNN and transfer learning with VGG16. These approaches provide a comprehensive analysis of COVID-19, Pneumonia, and Normal cases by leveraging various perspectives on the chest X-ray dataset.

- *Traditional Machine Learning Models*

The traditional machine learning models Naive Bayes, Decision Tree, K-Nearest Neighbors, and Support Vector Machine serve as baseline classifiers. Each image is reshaped into a two- dimensional format, where each pixel acts as a feature. These models contribute distinct classification approaches within the ensemble:

Naive Bayes assumes feature independence, enabling efficient computation. Decision Tree captures non-linear relationships, offering interpretable decision boundaries. KNN operates by clustering cases based on feature space proximity, while SVM maximizes separation between classes by identifying optimal decision boundaries. Though limited in spatial feature extraction, these models provide varied perspectives, establishing a foundational baseline for the ensemble.

- *Convolutional Neural Networks (CNN)*

A custom CNN model is designed specifically for the multi-class classification task of chest X- ray images. The CNN is trained on augmented data, which introduces variability, enabling the model to capture complex patterns in the images that may signify specific conditions.

The CNN is compiled with a low learning rate for stable training and categorical cross-entropy loss to manage the multi-class classification. Dropout layers are incorporated to prevent overfitting, enhancing the model's generalizability. The convolutional and pooling layers enable the CNN to capture spatial hierarchies within the data, from basic edges to intricate textures, making it particularly suited to identify patterns associated with COVID-19 and other cases.

- *Transfer Learning with VGG16*

Transfer learning is utilized with the VGG16 model, which has been pre-trained on the ImageNet dataset. The original classification layers are removed, and VGG16 is adapted with custom fully connected layers for the chest X-ray classification task.

The convolutional layers of VGG16 are frozen to preserve the learned feature extraction capabilities, such as edge detection and texture recognition, which generalize well to medical imaging. By adding custom layers, the model can focus specifically on distinguishing between COVID-19, Pneumonia, and Normal cases. This transfer learning approach leverages VGG16's pre-trained features, improving the model's generalization and boosting classification accuracy without requiring extensive training.

Integrating traditional models, the custom CNN, and transfer learning with VGG16 provides a diverse ensemble, enhancing the system's ability to robustly classify chest X-ray images.

➢ *Ensemble Method and Majority Voting*

The ensemble approach integrates predictions from all models using majority voting, wherein the predicted class is determined by the majority output across the models. This technique capitalizes on the diversity among classifiers, if one model misclassifies a sample, another may classify it correctly, enhancing overall accuracy. The ensemble method, by averaging multiple perspectives, seeks to yield a robust and reliable classification outcome, thus improving COVID-19 detection accuracy.

➢ *Evaluation Metrics and Visualization*

For model evaluation, metrics such as accuracy, classification reports (precision, recall, and F1- score), and confusion matrices are calculated for each classifier and the ensemble. These metrics help assess model performance comprehensively. To facilitate an intuitive comparison, model accuracies are visualized using a bar plot, offering a clear depiction of each model's effectiveness, with the ensemble approach expected to achieve the highest accuracy among them.

These implementation modules collectively form a robust ensemble-based framework for COVID-19 detection in chest X-rays. By leveraging both traditional and deep learning models, along with transfer learning and data augmentation techniques, the ensemble method maximizes predictive accuracy and offers a comprehensive diagnostic tool for enhanced COVID-19 and pneumonia detection.

V. RESULTS AND DISCUSSIONS

➢ *Image Preprocessing and Data Augmentation*

Images are loaded from folders based on class labels ("Covid," "Viral Pneumonia," and "Normal") and converted to RGB format for consistent color channels. Each image is resized to 128x128 pixels, ensuring a uniform input shape across the dataset. Pixel values are then normalized by dividing by 255.0, scaling them between 0 and 1 to improve model training.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

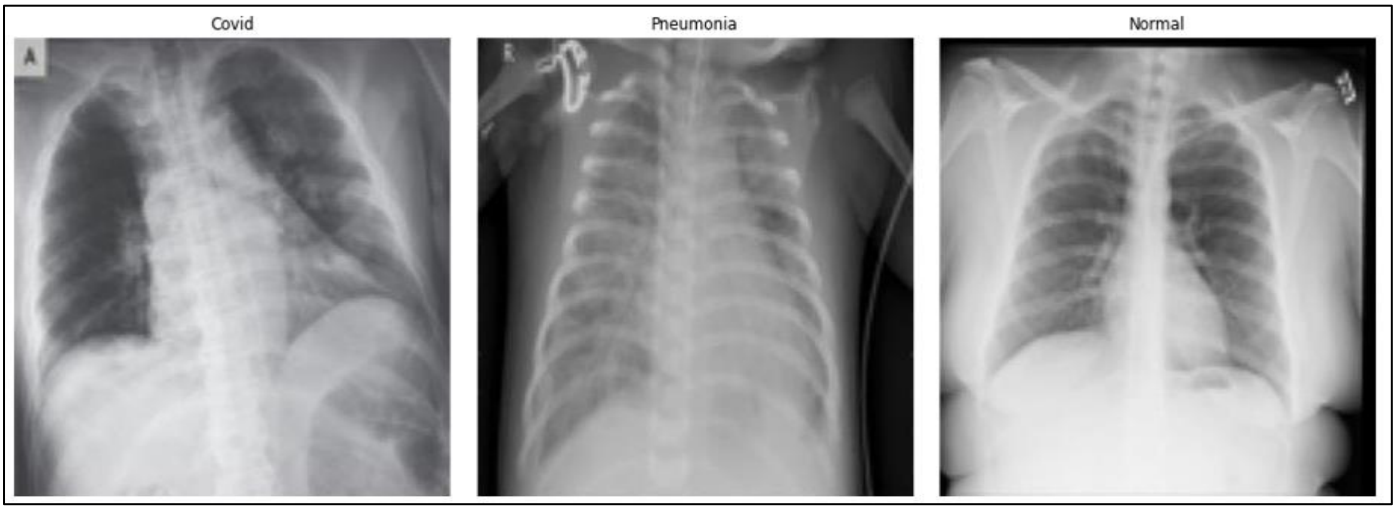

Fig 8 Pre-Processed Images from Each Label

Labels are encoded into numerical values to make them compatible with machine learning models. Data augmentation is applied to the training images, using techniques like rotation, zoom, and horizontal flips to create variations and prevent overfitting. For traditional machine learning models, images are flattened into 1D arrays to ensure compatibility. The augmented data generator is fitted to the training set for real-time augmentation during CNN training. This preprocessing pipeline ensures image quality, uniformity, and data diversity, preparing the dataset for effective model training and evaluation.

➢ *Calculated Metrics of Base Models*

The table summarizes classification performance metrics for six models Naive Bayes (NB), Decision Tree (DT), k-Nearest Neighbors (KNN),

Table 1 Calculated Metrics of Base Models

	PRECISION						RECALL					
	NB	DT	KNN	SVM	CNN	VGG16	NB	DT	KNN	SVM	CNN	VGG16
COVID	0.66	0.9	0.95	0.95	0.97	1	0.65	0.82	0.86	0.89	0.85	0.74
PNEUMONIA	0.66	0.93	0.89	0.95	0.94	0.98	0.91	0.98	0.99	0.99	0.97	0.91
NORMAL	0.58	0.86	0.9	0.91	0.87	0.77	0.49	0.9	0.89	0.94	0.95	0.99
MACRO AVG	0.63	0.9	0.91	0.94	0.93	0.92	0.64	0.9	0.91	0.94	0.92	0.88
WEIGHTED AVG	0.63	0.9	0.91	0.94	0.92	0.91	0.62	0.9	0.91	0.94	0.92	0.88

	F1 SCORE						SUPPORT
	NB	DT	KNN	SVM	CNN	VGG16	
COVID	0.66	0.86	0.9	0.92	0.91	0.94	930
PNEUMONIA	0.76	0.96	0.94	0.97	0.95	0.98	746
NORMAL	0.49	0.88	0.89	0.93	0.91	0.94	1003
MACRO AVG	0.64	0.9	0.91	0.94	0.92	0.95	2679
WEIGHTED AVG	0.62	0.9	0.91	0.94	0.92	0.95	2679
ACCURACY	0.64	0.9	0.91	0.94	0.92	0.95	2679

Support Vector Machine (SVM), Convolutional Neural Network (CNN), and VGG16—in detecting COVID-19, pneumonia, and normal cases from chest X-rays. Metrics include precision, recall, F1 score, and accuracy across three classes, with macro and weighted averages for overall performance.

VGG16 consistently demonstrates high precision, recall, and F1 scores, achieving the highest precision (1.0) in COVID-19 detection and strong scores for pneumonia (0.98) and normal cases. Its overall accuracy is the highest at 0.95, followed by CNN (0.92). In contrast, Naive Bayes has lower performance across all metrics, with an accuracy of 0.64, indicating its limited classification capability.

For recall, KNN, SVM, and CNN achieve near-perfect scores (0.99) for pneumonia, and VGG16 performs exceptionally well for normal cases with a recall of 0.99. KNN and SVM also show competitive F1 scores across all classes, although slightly lower than CNN and VGG16.

➢ *Confusion Matrix*

These confusion matrix heatmaps offer a clear view of the classification performance across three classes in a model, highlighting both correct predictions and misclassifications. Each matrix displays counts where each predicted class matches the actual class, with diagonal values (like 603, 682, and 423) representing true positives. These correct predictions showcase where the model performs well, giving insights into its strengths.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

Off-diagonal values illustrate the misclassifications, where predictions deviate from the actual class labels. For example, a value of 30 in one matrix shows instances where class 0 was mistaken for class 1, while a higher count, like 230, highlights a more significant error where class 0 was predicted as class 2. This error pattern can help identify classes that are frequently confused, which may indicate overlapping features or areas needing more data.

The color intensity in each matrix visually reinforces these insights, with darker shades indicating higher counts. This gradient effect allows for a quick glance assessment of which errors are more prevalent and which classes achieve high accuracy. The confusion matrix, therefore, not only quantifies model performance but also pinpoints areas for potential improvement, particularly where misclassifications are frequent and dominant.

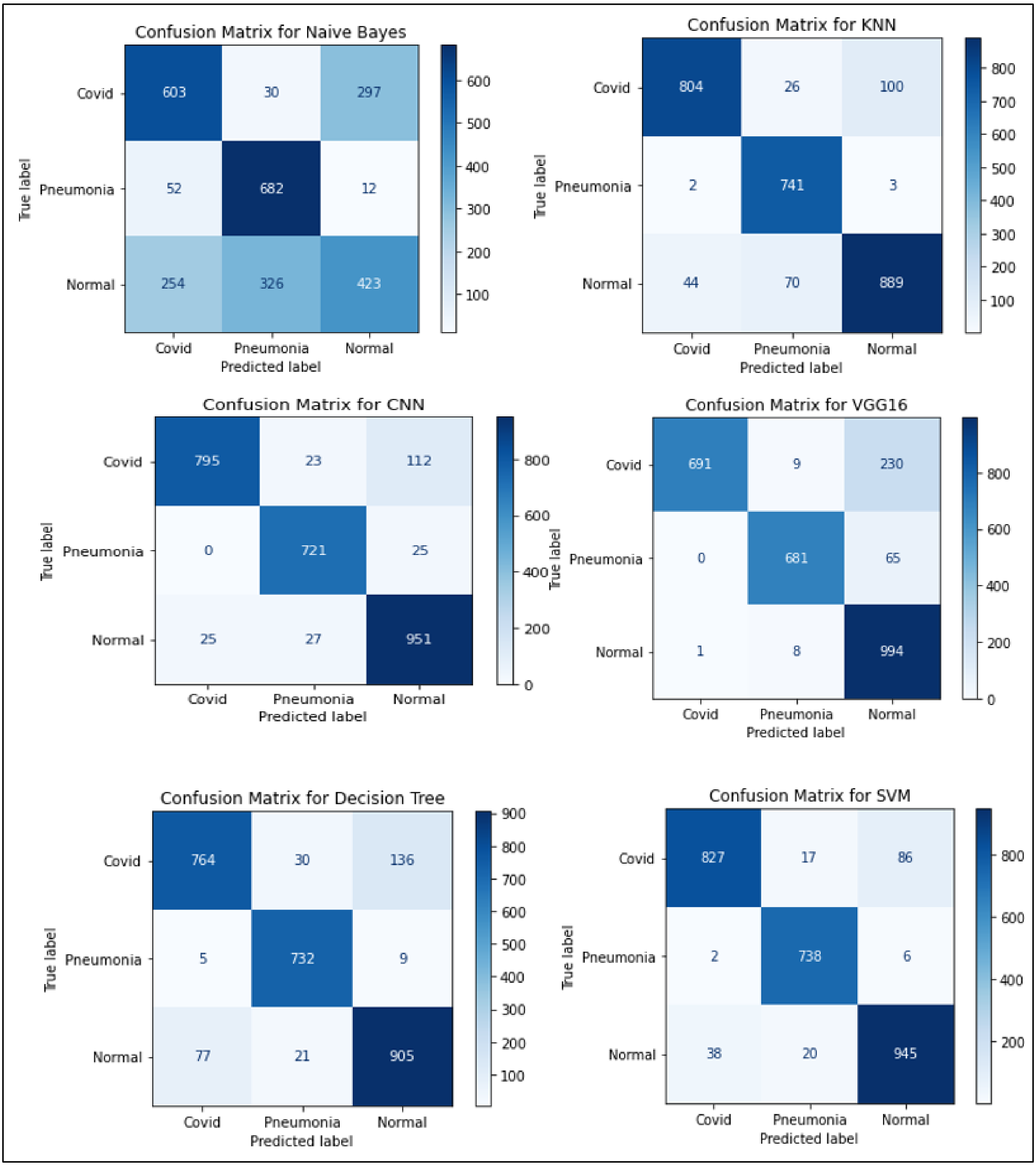

Fig 9 Confusion Matrices of six models

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

➢ *Training and Accuracy Loss and Accuracy*

In machine learning, especially in training Convolutional Neural Networks (CNNs) like VGG16, metrics such as training loss and accuracy are critical for evaluating model performance. Training loss is a measure of how well the model's predictions align with the actual labels in the training data, indicating how much error the model is making. A lower loss suggests better predictions. Training accuracy, on the other hand, indicates the percentage of correctly classified samples in the training dataset. Monitoring both metrics helps identify how well the model is learning patterns in the data and if it is overfitting or underfitting.

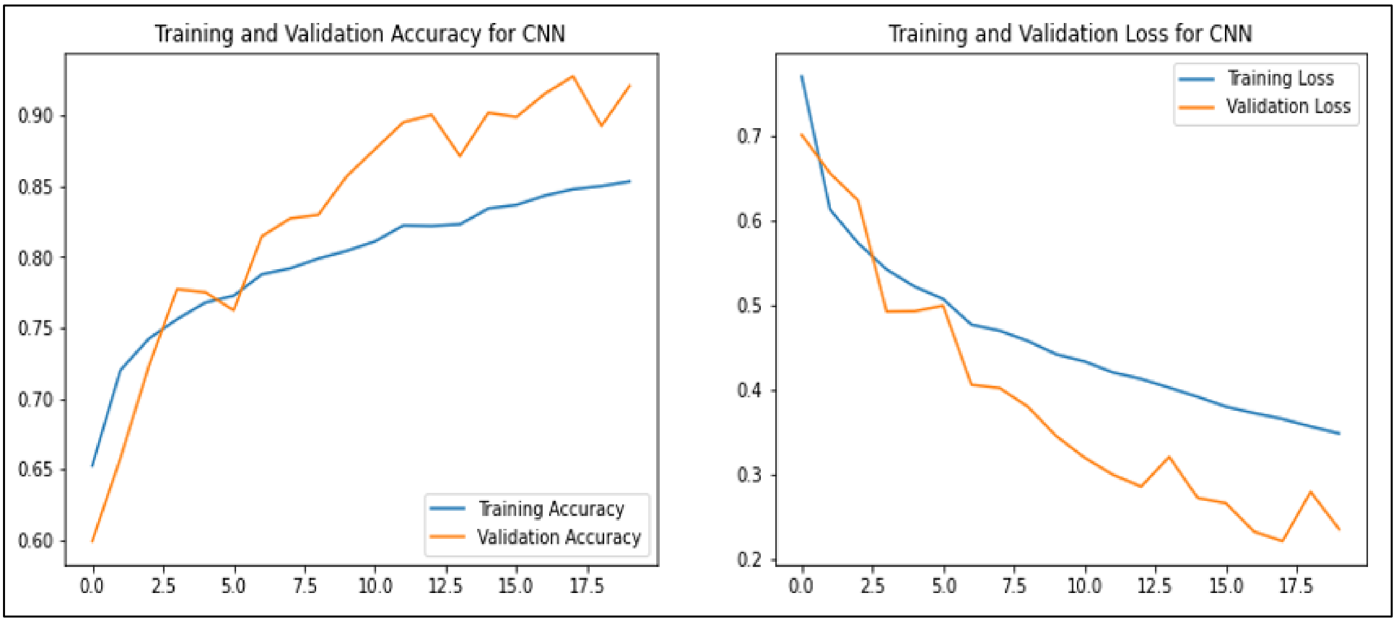

Fig 10 Training and Validation Accuracy and Loss for CNN

CNNs are widely used for image-based tasks due to their ability to automatically learn spatial hierarchies of features. VGG16, a popular CNN architecture, is known for its deep layers and consistent performance in image recognition tasks. VGG16 employs a series of convolutional and pooling layers, followed by fully connected layers, which capture increasingly complex patterns in the data. By monitoring loss and accuracy curves during training, practitioners can fine-tune the model to achieve better generalization and avoid overfitting.

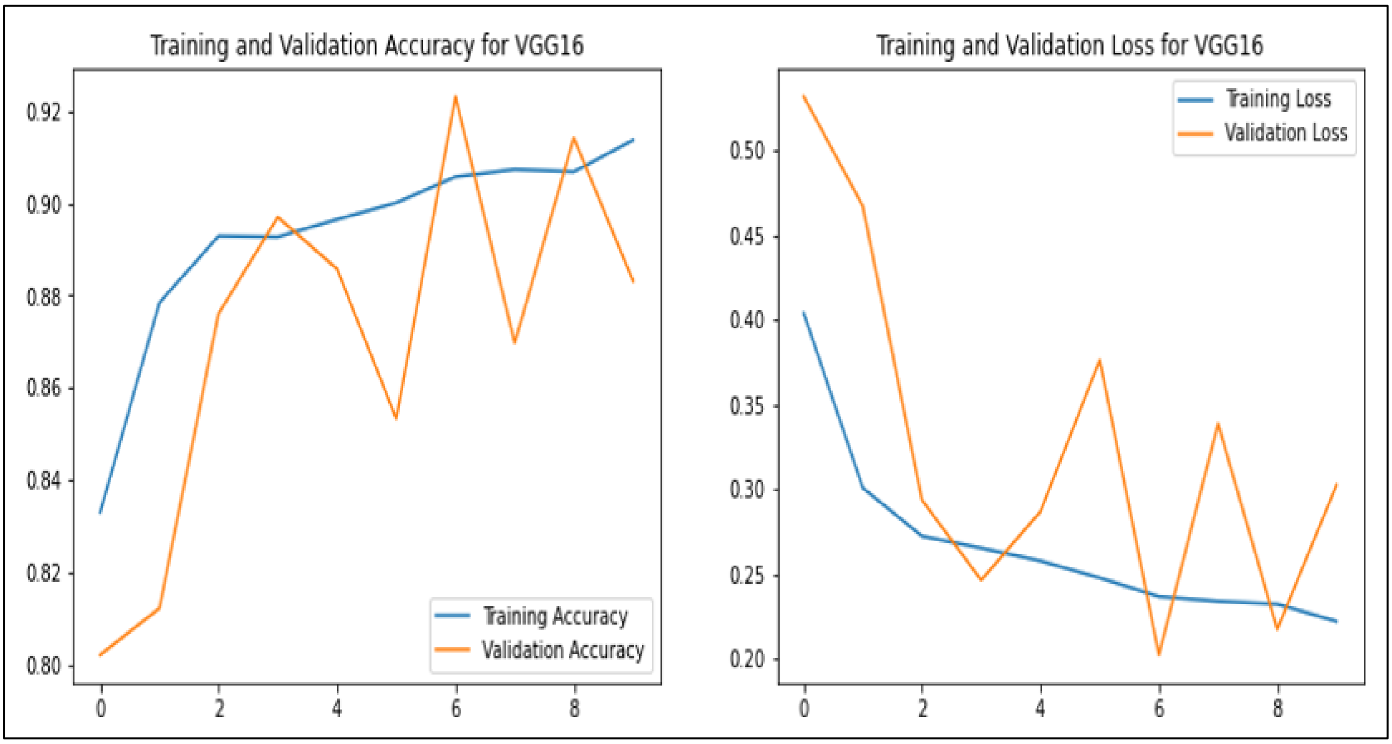

Fig 11 Training and Validation Accuracy and Loss for VGG16

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

The figure displays two-line graphs representing the training and validation performance metrics of a model across epochs.

The Left Graph shows training accuracy and validation accuracy over time. The training accuracy steadily increases, while the validation accuracy fluctuates more. The variations in the validation accuracy line may suggest some instability in the model's performance on unseen data, which could indicate mild overfitting.

The Right Graph shows training loss and validation loss. The training loss consistently decreases, as expected during training. However, the validation loss fluctuates significantly, suggesting that the model's generalization on new data is inconsistent.

➢ *Ensemble Model*

To create the ensemble model, the combined predictions from various machine learning and deep learning models, specifically Decision Tree (DT), Naive Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Convolutional Neural Network (CNN), and VGG16. Each of these models was trained on the same dataset and contributed to the ensemble through a majority voting mechanism. This approach leverages the strengths of each model and reduces the likelihood of overfitting or bias from any single model.

Table 2 Confusion Matrix and Other Metrics for Ensemble Model

	PRECISION	RECALL	F1-SCORE	SUPPORT
COVID	0.96	0.93	0.94	930
PNEUMONIA	0.96	1	0.98	746
NORMAL	0.94	0.94	0.94	1003
MACRO AVG	0.95	0.96	0.95	2679
WEIGHTED AVG	0.95	0.95	0.95	2679
ACCURACY			0.95	2679

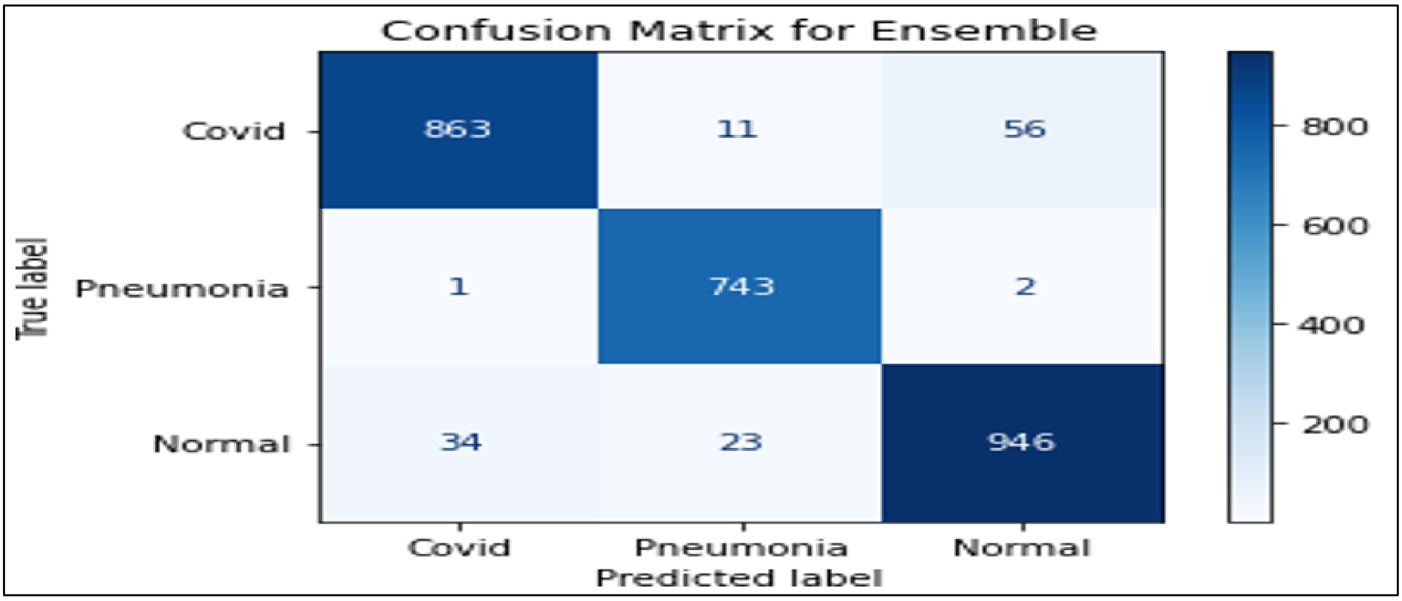

Fig 12 Confusion Matrix and Other Metrics for Ensemble Model

The confusion matrix and classification report provide insights into the ensemble model's performance across three classes: COVID, Pneumonia, and Normal.

The confusion matrix analysis shows the model performs well across all classes with a high number of correct predictions. Specifically, the model accurately predicted 863 COVID cases with some misclassifications, nearly perfect results for Pneumonia with 743 correct predictions, and 946 correct predictions for Normal cases.

The classification report metrics indicate strong performance. The model achieved a precision of 0.96, recall of 0.93, and an F1-score of 0.94 for COVID, while Pneumonia had outstanding results with a 0.96 precision, perfect recall of 1.0, and an F1-score of 0.98. Normal cases also showed robust performance with scores of 0.94 for precision, recall, and F1-score.

Both macro and weighted averages for precision, recall, and F1-score are 0.95, indicating balanced performance across all classes. The overall accuracy of 0.95 demonstrates the model's reliability in classifying 95% of the instances correctly.

In summary, the ensemble model is highly effective for distinguishing between COVID, Pneumonia, and Normal cases, with high precision, recall, and F1-scores. This makes the model dependable for medical diagnosis, especially in identifying Pneumonia cases accurately.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

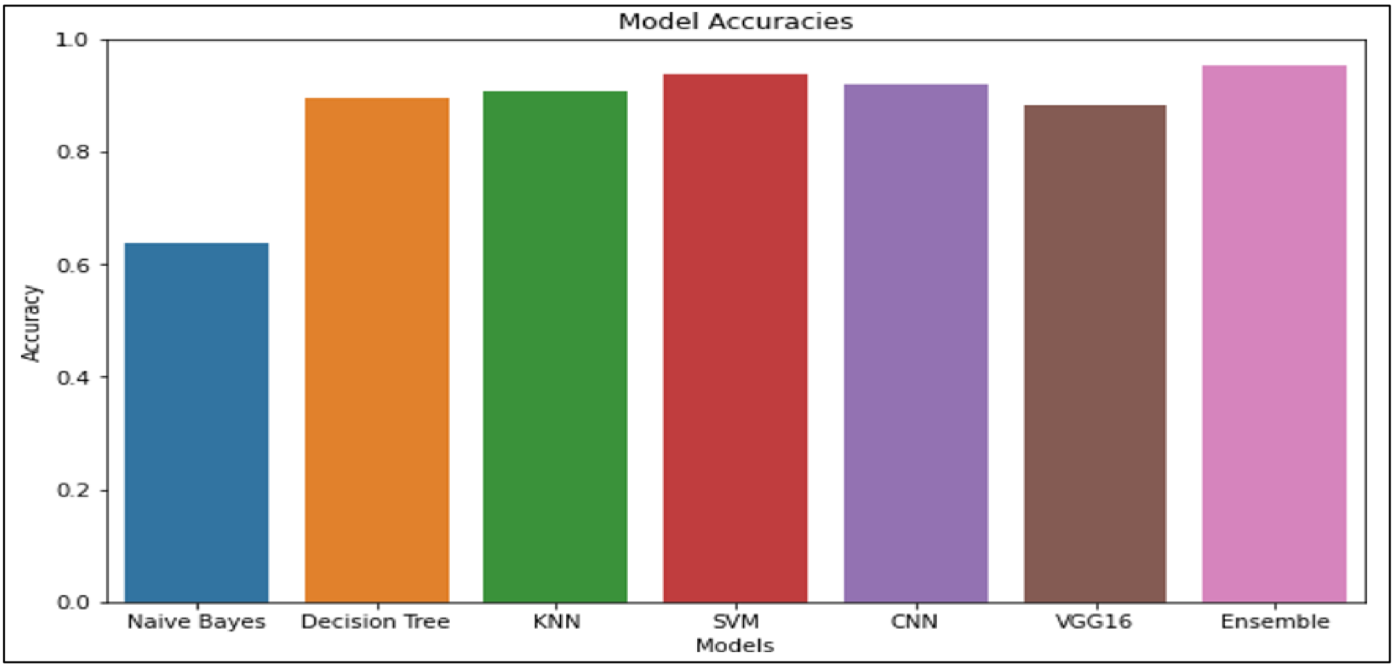

Fig 13 Comparison of Model Accuracies with Ensemble Model

The Bar Graph, titled "Model Accuracies," illustrates the performance of various machine learning models in classifying data, with a focus on their accuracy. The x-axis lists seven different models: Naive Bayes, Decision Tree, KNN, SVM, CNN, VGG16, and an Ensemble model. Each bar represents the accuracy of one model, with values on the y-axis ranging from 0.0 to 1.0, indicating the proportion of correct predictions made by each model.

The chart shows that Naive Bayes has the lowest accuracy among the models, scoring below 0.7. In contrast, Decision Tree, KNN, SVM, CNN, and VGG16 models achieve high accuracies, all close to or above 0.9. The Ensemble model, which combines predictions from multiple models using majority voting, outperforms the others, achieving the highest accuracy of all the methods tested.

This comparison suggests that while individual models can perform well, using an Ensemble model can provide even better accuracy. Ensemble techniques often improve performance by leveraging the strengths of multiple models, reducing errors, and enhancing prediction reliability. This finding indicates the value of ensemble methods in achieving higher predictive accuracy, particularly in complex classification tasks.

- *Implemention of Ensemble Model in an Application*

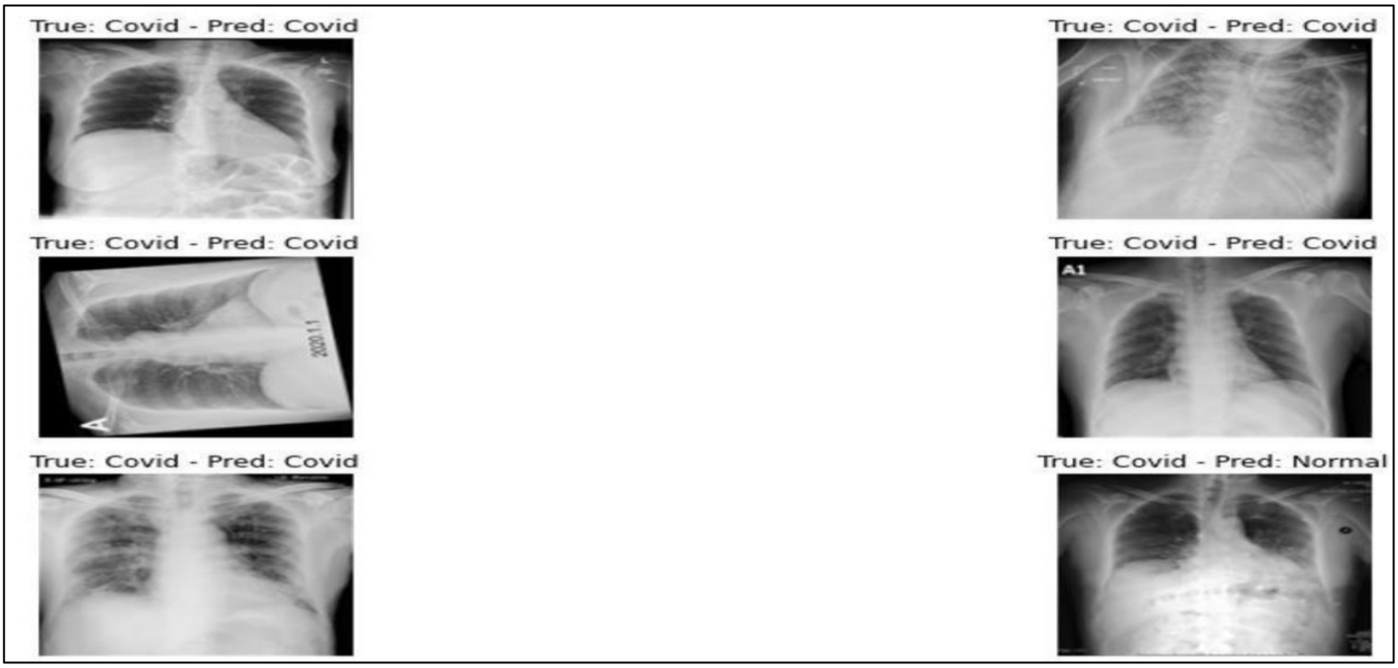

Fig 14 Few Predictions Made Using the Ensemble Model

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

The ensemble model consisting of multiple machine learning models to predict COVID-19, Pneumonia, and Normal cases from chest X-ray images is implemented in an app. This app allows users to upload their X-ray images, which are then processed and analyzed by a variety of models such as CNNs, VGG16, Naive Bayes, Decision Tree, KNN, and SVM. By combining the strengths of these models, the app ensures high accuracy and reliable predictions. The ensemble approach helps in mitigating individual model limitations, providing a balanced and effective diagnostic tool. This app aims to aid in efficient COVID-19 triage and patient care, making it a valuable tool in the healthcare sector.

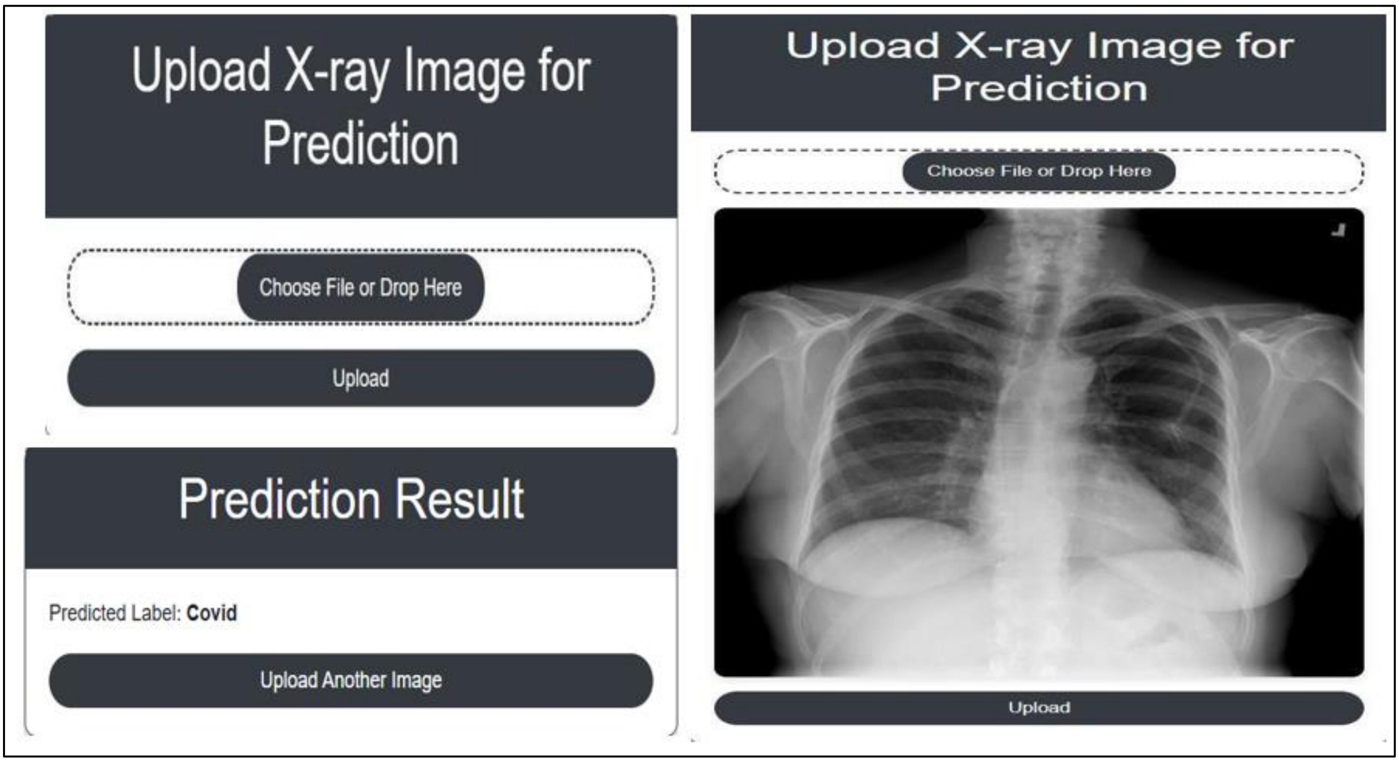

Fig 15 Application Implementation of Ensemble Model

VI. CONCLUSION AND FUTURE WORK

Ensemble-based approach for detecting COVID-19 in chest X-ray images, offering a robust, accurate, and reliable tool for automated medical diagnostics. By utilizing a majority-vote ensemble of five distinct supervised algorithms, the model benefits from diverse classifier perspectives, enhancing diagnostic precision and consistency. Advanced feature extraction techniques, such as statistical texture descriptors, allow the model to detect specific radiographic patterns associated with COVID-19. The addition of data augmentation further strengthens the model's ability to adapt across varied datasets, ensuring stable performance in different clinical settings. This approach meets the urgent need for efficient COVID-19 screening and highlights machine learning's transformative role in healthcare, especially in improving diagnostic speed and accuracy.

Future developments could extend this model's capability to diagnose other infectious and respiratory diseases by adapting it for different imaging data types, like CT scans, or incorporating clinical information. Employing advanced augmentation techniques, such as GANs, may further enhance model robustness for diverse imaging conditions. Integrating this system into real-time clinical workflows could provide rapid decision support to healthcare professionals, enhancing its practical value. Additionally, improving model interpretability would enable clinicians to understand crucial diagnostic features, fostering trust in AI-assisted tools. This research lays a foundation for advancing machine learning in healthcare, emphasizing AI's potential for effective and timely infectious disease detection.

REFERENCES

[1]. Aleka Melese Ayalew, Ayodeji Olalekan Salau, Yibeltal Tamyalew, Bekalu Tadele Abeje, Nigus Woreta, "X-Ray image-based COVID-19 detection using deep learning", In: Multimedia Tools and Applications 82, Springer, 2023, vol: 82, pp: 44507-44525.
[2]. Sreeparna Das, Ishan Ayus, Deepak Gupta, "A comprehensive review of COVID-19 detection with machine learning and deep learning techniques", In: Health and Technology 13, springer, 2023, vol:13, pp:679-692.
[3]. Lin Zou, Han Leong Goh, Charlene Jin Yee Liew, Jessica Lishan Quah, Gary Tianyu Gu, Jun Jie Chew, Mukundaram Prem Kumar, Christine Gia Lee Ang, and Andy Wee An Ta, "Ensemble Image Explainable AI (XAI) Algorithm for Severe Community-Acquired pneumonia and COVID-19 Respiratory Infections", In: IEEE transactions on Artificial Intelligence, IEEE,

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26feb917

2023, vol:4, pp:242-254.
[4]. Haval I. Hussein, Abdulhakeem O. Mohammed, Masoud M. Hassan, Ramadhan J. Mstafa, "Lightweight deep CNN-based models for early detection of COVID-19 patients from chest X-ray images", In: Expert Systems with Applications, Elsevier, 2023, vol:223.
[5]. Vandana Bhattacharjee, Ankita Priya, Ankita Priya, Shamama Anwar, "DeepCOVNet Model for COVID-19 Detection Using Chest X-Ray Images", In: Wireless Personal Communications 130, Springer, 2023, vol:130, pp: 1399-1416.
[6]. Truong Dang, Truong Dang, John McCall, Eyad Elyan, Carlos Francisco Moreno-García, "Two- layer Ensemble of Deep Learning Models for Medical Image Segmentation", In: Cognitive Computation 16, Springer, 2024, vol: 16, pp:1141-1160.
[7]. Talib Iqball, M. Arif Wani, "Weighted ensemble model for image classification", In: Int. j. inf. Tecnol. 15(2), Springer, 2023, vol: 15, pp:557-564.
[8]. Baijnath Kaushik, Akshma Chadha, Reya Sharma, "Performance Evaluation of Learning Models for the Prognosis of COVID-19", In: New Generation Computing 41, Springer, 2023, vol: 41, pp: 533-551.