PneuML: A NOVEL SEQUENTIAL CONVOLUTIONAL NEURAL NETWORK-BASED X-RAY DIAGNOSTIC SYSTEM FOR PNEUMONIA IN CONTRAST TO MACHINE LEARNING AND PRE-TRAINED NETWORKS

Sunil KUMAR[1,2], Harish KUMAR[1]

Pneumonia is a formidable global health challenge, contributing significantly to worldwide mortality rates. This research addresses this issue by developing and evaluating a sophisticated image classification model, PneuML, to accurately predict pneumonia cases using chest X-ray images. Machine learning (ML) techniques, pre-trained convolutional neural networks (CNNs), and a novel PneuML sequential CNN architecture were rigorously examined to achieve this objective. The study's primary focus was to assess the effectiveness of the PneuML model in classifying X-rays. The results obtained from the PneuML model were highly promising, achieving an exceptional accuracy rate of 96.21% and an impressive F1 score of 95.65%. These outcomes surpassed the performance of both pre-trained CNNs, with ResNet50 as the top-performing model, and traditional ML methods, where XgBoost ML classifiers were best performed. This research also facilitated a comparative analysis between CNNs and ML, underscoring the superiority of the PneuML CNN architecture in pneumonia classification through X-ray imaging. Furthermore, the study explored advanced techniques such as class weights to enhance model performance. Incorporating these cutting-edge concepts holds promise for further refinement and improved results.

Keywords: CNN, DenseNet121, ResNet50, SVM, VGG16, XgBoost

1. Introduction

The illness produces discomfort or incapacity due to an internal malfunction caused by lung disease. Individuals may be unaware of lung diseases or conditions since diagnosis is difficult. This is because most lung diseases initially do not have symptoms before they produce discomfort. We can cure lung disease if we diagnose it. Traditionally, medical professionals have manually diagnosed lung diseases by utilizing procedures grounded in the symptoms of the conditions. It is a common belief that the most lethal diseases, such as brain or blood cancer, result in the most

[1] Research Scholar, Department of Computer Engineering, J.C. Bose University of Science and Technology, YMCA, Faridabad, India. e-mail: sunilymca1@jcboseust.ac.in
[1] Associate Professor, Department of Computer Engineering, J.C. Bose University of Science and Technology, YMCA, Faridabad, India. e-mail: htanwar@jcboseust.ac.in
[2] Assistant Professor, Department of Information Technology, School of Engineering and Technology (UIET), CSJM University, Kanpur, India.

significant number of fatalities worldwide. This belief, however, is inappropriate because several of these diseases do not rank among the top 10 global causes of mortality. Unexpectedly, lung disorders account for the majority of fatal diseases. The WHO identified the ten deadliest diseases that claimed the most lives worldwide between 2000 and 2019 [1].

Lower respiratory infections are the most dangerous contagious illness globally and the fourth most common reason for death. Lower respiratory disorders encompass a variety of diseases that may affect the lungs, the most well-known of which is pneumonitis [2]. When the lungs get infected with pneumonia, it can lead to swelling and fluid accumulation. Many different kinds of microorganisms are capable of causing it, including bacteria, viruses, and fungi. The seriousness of pneumonia may range from mild to potentially fatal, and treatment depends on the infection's source and severity. There are several types of pneumonia, each with different causes and treatment options. Some common types of pneumonia include:

• Viral pneumonia: Pneumonia caused by the virus is less common but can be severe; one example is COVID-19.

• Fungal pneumonia: A fungus causes this type of pneumonia. People with an impaired immune system are more likely to be affected.

• Mycoplasma pneumonia: This type of bacteria called Mycoplasma pneumonia causes this type of pneumonia, often referred to as 'walking pneumonia,' as it can induce mild symptoms resembling a cold or the flu.

• Community-acquired pneumonia: Bacteria or viruses commonly present in the community typically cause this.

• Hospital-acquired pneumonia: An individual can only catch this particular strain if they are a patient in a public hospital. Antibiotic-resistant bacteria typically cause it [3].

Fig. 1 displays X-ray illustrations of pneumonia and normal, obtained from the Large Dataset of Labelled Optical Coherence Tomography and Chest X-Ray (LDOCTCXR) dataset available to the public [4].

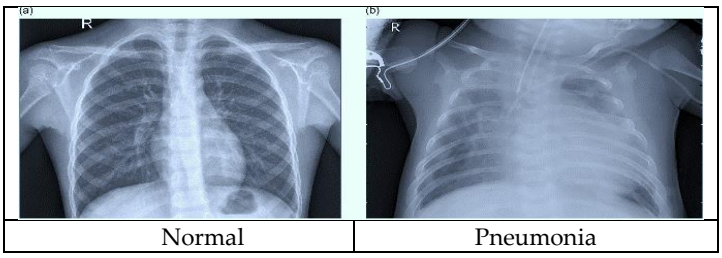

Fig. 1. Instances of chest X-ray from LDOCTCXR dataset [4]

X-rays are the imaging modalities that are employed most frequently in the process of diagnosing pneumonia. They can show characteristic signs of pneumonia, such as inflammation and fluid accumulation in the lungs. Interdisciplinary fields are combined in the new era to provide new real-life problem solutions.

ML in medical services is a recent trend that offers novel approaches to real-world challenges. ML is an artificial intelligence (AI) that can analyze medical data and make predictions. In the context of pneumonia diagnosis, ML algorithms can analyze X-ray images and predict whether or not a person has pneumonia [5]. ML algorithms offer promising solutions for the task. This research presents a comprehensive approach to pneumonia diagnosis using ML techniques and feature extraction. The first step extracts relevant features using the local binary pattern (LBP) method [6]. Subsequently, the supervised learning techniques train the ML model, including Support-Vector-Machine (SVM) [7], K-nearest-neighbor (KNN) [8], random forest (RF) [9], gradient boost [10], Extreme Gradient Boosting (Xgboost) [11], and Adaptive Boosting (AdaBoost) [12]. Successful validation could lead to the deployment of the ML model. However, it is essential to note that ML models need proof on a large and diverse X-ray dataset before being used in the clinical setting [13].

One common approach is using deep learning algorithms, such as CNNs, to analyze X-ray scans for signs of pneumonia. CNNs are a specific kind of ML algorithm that works very well when used for the problem of image analysis. CNNs autonomously extract the most salient or intricate features and patterns from the given input X-ray images through training. Various architectures, including VGG16, VGG19 [14], ResNet50 [15], InceptionV3 [16], and DenseNet121 [17], were employed to train the CNN models in the investigation. Utilizing ML and CNN algorithms on extensive labeled image datasets to analyze pneumonia presence or absence indicators. Once trained, the model can be employed to investigate new images to make a diagnosis [18].

The motivation of this investigation is to address the considerable global health issue of pneumonia, which significantly contributes to mortality rates worldwide, and to create and assess an effective image classification model specially tailored to effectively forecast instances of pneumonia by using X-ray images. The proposed PneuML architecture, meticulously designed and effectively demonstrated in the research, employed a 12-layer sequential CNN model. The architecture comprises a solitary input layer, three convolutional layers, three maximum pooling layers, one flattening layer, two dense layers, and a dropout layer. The primary contributions of this research are outlined below:

• Pneumonia is a substantial worldwide health concern, leading to elevated rates of fatalities and addressing a global health challenge. This investigation aims

to tackle this problem by constructing an advanced PneuML image classification approach.

• The investigation suggested that it collected pneumonia and normal images by merging two publicly available datasets to create a novel dataset.

• The PneuML sequential CNN architecture performed better than pre-trained CNNs and ML classifiers.

• The investigation investigated sophisticated methodologies, such as the utilization of class weights, to augment the performance of the existing and recommended PneuML model.

Furthermore, the following outline depicts the article's organization: Section 2 presents the literature review. Section 3 addresses the materials and methods being used to perform the investigation. Section 4 offers the result and discussion section, ML, CNN, and PneuML sequential CNN architecture findings. Section 5 concludes the research.

2. Literature Review

X-rays were the subject of an investigation by researchers who examined various diagnostic approaches for pneumonia in published studies. Two distinct methods, traditional ML and CNN, were observed when analyzing the articles.

To train and verify deep CNNs, transfer learning with image augmentation was employed [5]. The convolutional network's distinguishing feature was the use of dropout. Kaggle's diagnostic imaging competition provided 5856 tagged images for training and testing, where the network achieved 97.2 % accuracy, 97.3% recall, 97.4% precision, and an AUC of 98.20% [19]. Twelve pre-trained CNN models were fine-tuned to differentiate healthy X-rays from pneumonia cases using 6,330 images for training, validation, and testing. Most architectures performed well, with a mean f1-score of 84.46% in distinguishing the four classes [20]. A multi-branch fusion auxiliary learning (MBFAL) technique achieved a classification performance of 95.61% for pneumonia detection using lung X-rays. MBFAL demonstrated a standard accuracy of 98.70%, 99.10%, 96.60%, and 96.80% for multi-classification, with recalls of 97.20%, 98.60%, 96.10%, and 89.20% [21]. The paper presented a "CJT-based ensemble of classifiers," demonstrating the superiority of Condorcet's jury theorem-based ensembles of learners regarding accuracy. The ensemble approach combined CNN model findings to enhance COVID-19 pneumonia detection in CXR images [22]. Researchers employed novel imaging techniques for pneumonia identification and tracking using thermography [23]. The appropriate CXR image features were chosen using a hybrid social group optimization algorithm, and support vector classifiers achieved good classification results [24]. The presented method, Multi-Scale Attention Network (MSANet), utilized deep learning for classifying COVID-19 and other pneumonia forms with

improved accuracy through automated consideration of discriminatory data and multi-scale features [25,26]. The approach presented as SVD-CLAHE Boosting, ResNet-50, and BWCCE showcased remarkable results, achieving an accuracy of 94.00% and an F1 score of 95.00%. This approach uniquely combined techniques such as Singular Value Decomposition (SVD) and Contrast Limited Adaptive Histogram Equalization (CLAHE) in conjunction with a deep learning architecture, ResNet-50, and Balanced Weighted Cross-Entropy (BWCCE) loss functions. The success of this approach underscored the potential of combining multiple techniques for robust and precise results [27]. In the context of transfer learning, the Transfer Learning with Inception V3 approach displayed strong performance, having achieved an accuracy of 92.80%. Transfer learning, specifically through the Inception V3 model, highlighted the value of leveraging pre-trained models to achieve accurate predictions [28]. Furthermore, the "Mobile Freeze-Net" method showcased an accuracy of 93.00% and an F1 score of 94.00%, illustrating the efficacy of mobile-oriented architectures [29]. Additionally, adopting more traditional methodologies, such as utilizing a CNN, resulted in an accuracy of 93.75% and an F1 score of 94.00% [30]. The exploration confirmed the reliability of the established machine and deep-learning techniques in tasks. Exploring these various approaches enhanced our understanding of the strategies employed in the job.

3. Material and Methods

The presentation of the fundamental architecture of our PneuML intended systematic perspective in contrast to ML and pre-trained CNN (Fig. 2). The system consists of multiple processes, including the collection of lung X-ray image datasets, the preprocessing and extracting of the necessary features, the construction of classification models based on ML and CNN and training them, and their evaluation based on specific performance metrics.

3.1 X-Ray Image Dataset

The investigation involved pneumonia and normal images from a generated dataset by combining two source datasets. The first dataset, LDOCTCXR, encompassed the acquisition of anterior-posterior X-ray images specifically targeting pediatric children aged one to five [4].

In contrast, the second dataset was called the Balanced Augmented COVID CXR dataset [31]. The Balanced Augmented COVID CXR dataset had four classifications: COVID, Normal, Lung Opacity, and Pneumonia. In our investigation, we selected only pneumonia images and normal classes. The methodologies used in the second dataset included the utilization of under-sampling, over-sampling, and image-processing methods for data augmentation. The study employed a distinct approach known as SVD to generate images that

exhibited nuanced differences in brightness and contrast. CLAHE was applied to enhance the features of the images [27].

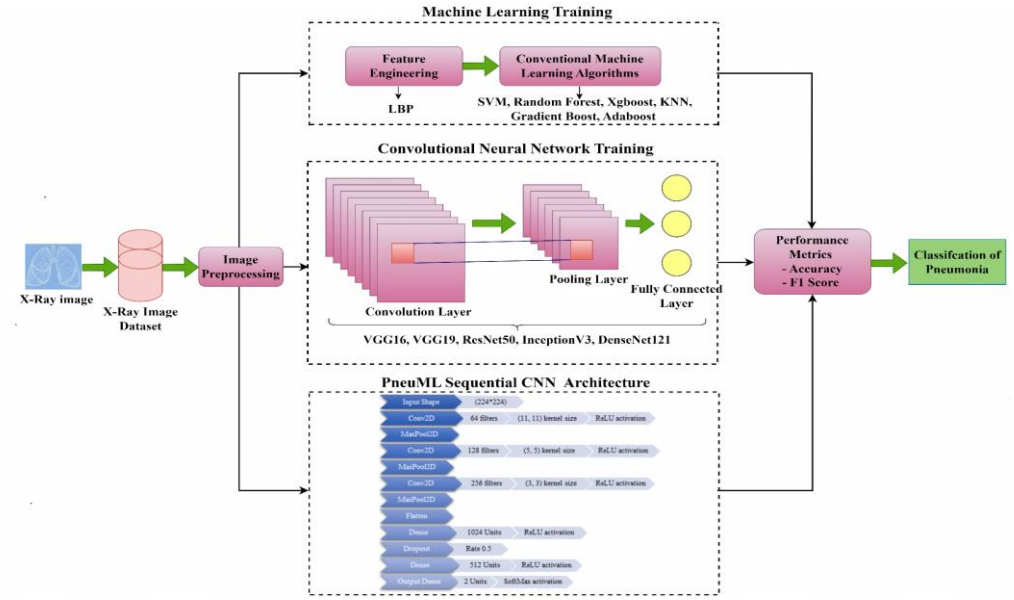

Fig. 2. Framework of PneuML systemic approach in contrast to ML and pre-trained CNN

The investigation revealed the benefits of adding more images and explained the reasoning behind the image selection. The created dataset is designed to enhance diversity and comprehensiveness by integrating two source datasets. These choices contributed to a more extensive and detailed dataset, which aided in detailing and improving the effectiveness of the investigation. Table 1 provided a comprehensive summary of the datasets, where the created dataset consisted of 19,435 samples. The generated dataset comprised 9,638 samples of pneumonia and 9,797 samples of normal lung conditions, resulting in a balanced dataset.

Table 1

Generated Dataset

Dataset	Pneumonia Samples	Normal Samples	Total
LDOCTCXR [4]	4,273	1,583	5,856
Balanced Augmented Covid CXR [31]	5,365	8,214	13,579
Generated	9,638	9,797	19,435

The investigation employed 80 percent to train and 20 percent to test the generated dataset. To eliminate biases from the dataset was randomized in preparation. It is important to note that the eminence of the image dataset is crucial

for the model's effectiveness. The dataset was balanced, meaning it should have a similar number of images for normal and pneumonia cases.

3.2 Image Pre-processing

The preprocessing tasks encompassed primary operations, including image resizing and dimension reduction. These techniques were deemed necessary due to discrepancies found in the dimensions of the data. It's worth highlighting that resizing and dimension reduction emerged as the two pivotal approaches for achieving effective preprocessing.

Some significant concerns came to light as we delved into the initial datasets. The first crucial step involved resizing images to a standardized dimension of 224 by 224 pixels. This decision was prompted by varied image sizes within the raw dataset. Ensuring uniformity across the training of the ML, CNN, and PneuML models hinged on maintaining consistent image dimensions. As a result, we immediately implemented a scaling procedure to align the dimensions. The samples initially were two-dimensional representations derived from three-dimensional source data [32].

3.3 ML Training

One potential use of ML was building a model using a dataset of X-ray images that included both pneumonia-positive and pneumonia-negative cases. Once the model was trained, it could analyze new X-ray images and predict whether a person had pneumonia. A stepwise ML approach to pneumonia detection typically involves the following steps:

• **Step-1**: **Feature extraction and selection:** The ML model was employed to extract relevant features that could be used to distinguish between healthy and infected lungs. The LBP, which by default contained 256 bins, was utilized in feature extraction. LBP considers an image's local texture information by comparing a pixel's intensity with its neighboring pixels. The Equation 1 of LBP was given as follows:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} 2^n * g\big(I(x_n, y_n) - I(x_c, y_c)\big) \quad (1)$$

Where x_c, y_c is the coordinates of the center pixel, x_n, y_n are the coordinates of the neighboring pixels, *I(x,y)* is the intensity of the pixel at axes *(x,y)*, and *g(x)* is a function defined as:

$$g(x) = 1, \text{ if } x >= 0$$
$$= 0, \text{ if } x < 0$$

The equation computed the LBP for a pixel in an image, where each bit represented the comparison of the intensity of the center pixel with that of its

neighbors. It showed that the LBP encoded the local texture around the pixel. Due to its effectiveness, the LBP operator was widely used in computer vision applications, such as texture classification [6].

• **Step-2: Model training:** After extracting and selecting the features, we provide the labelled dataset to the ML model for training. The ML model then learns the correlation between the features and the presence or absence of pneumonia. This process usually involves using classification algorithms within ML.

• **Step-3: Conventional ML Algorithms:** Our research used conventional ML algorithms, including SVM, KNN, RF, gradient boost, Xgboost, and AdaBoost. Each of these algorithms is described briefly below:

- SVM is a supervised learning method for classification and regression analysis. It determines optimal boundaries to separate data into different classes.

- KNN performs classification or regression using an instantiation-based, non-parametric learning strategy, which makes decisions based on most of its k-nearest neighbors' input.

- Random Forest executes both classification and prediction tasks. It generates multiple decision trees and aggregates their predictions to produce an outcome.

- Ensemble learning methods like gradient boosting leverage multiple weak learners to construct a single robust learner. This approach enables subsequent models to build upon the successes of their predecessors.

- XGBoost represents an advanced gradient-boosting technique. It efficiently handles large-scale and high-dimensional datasets.

- AdaBoost, an ensemble technique, addresses prior model errors by assigning greater weight to misclassified instances. It combines multiple weak learners to form a powerful learner [33].

• **Step-4: Model evaluation:** The trained model is then evaluated on a separate X-ray test dataset to assess its effectiveness. Typically, research accomplished this by using metrics accuracy and F1 score.

This stepwise approach is employed to train the ML models. However, it is imperative to highlight that the model's success will depend on the quality of the image used to teach it and the choice of ML algorithms.

3.4 CNN Training

CNNs are reliable, efficient, and highly effective ML techniques used for image analysis operations, like analyzing X-rays for diagnosing pneumonia [13]. Employing a CNN for the task includes the following steps, while other steps are the same as ML training:

• **Feature extraction, selection and CNNs Model training:** The explicitly selected and retrieved features are the foundation for implementing conventional ML techniques. Conversely, CNNs autonomously learn to extract the most

significant or intricate features and patterns from X-ray images. The CNN learns to identify patterns in images indicative of pneumonia presence [18]. After training the CNN with the designated X-ray dataset, it can analyze X-ray images and predict pneumonia in patients. Different architectures, such as VGG16, VGG19, ResNet50, InceptionV3, and DenseNet121, were employed to train the models. These models were all pre-trained using the ImageNet database [19].

Traditionally, Convolutional layers gathered detailed characteristics from input data in the experiment, allowing the network to discover patterns and correlations. Subsequent max-pooling techniques downscaled feature maps, keeping relevant information while lowering computing costs. The ReLU activation function introduced non-linearity, promoting complex representations necessary for the network's successful generalization.

- **Convolution Operation:**

The convolution process entails the movement of a filter, also known as a kernel, through an input data matrix. This movement comprises element-wise multiplication and subsequent summation of the overlapping sections, resulting in the extraction of features. Capturing spatial patterns within the data allows the neural network to identify and discern edges, textures, and other local features from the X-ray image. The Equation 2 of the convolution operation in a CNN is as follows:

$$(f * g)(x, y) = \sum\sum f(i, j) * g(x - i, y - j) \quad (2)$$

This equation calculates the dot product between the kernel and the corresponding region of the input feature map, summed over all positions *(i, j)* within the kernel. Where:

- *(f * g) (x, y)* represents the output value at axes *(x, y)* in the resulting feature map.
- *f(i, j)* represents the input value at position *(i, j)* in the input feature map.
- *g(x - i, y - j)* represents the kernel value at position *(x – i, y - j)* in the filter/kernel [34].

- **Max Pooling Operation:**

The process of max pooling involves the reduction of spatial dimensions in data by picking the maximum value from a collection of surrounding items within each area. The Equation 3 of the maximum pooling operation in a CNN is as follows:

$$\max\_pool(x, y) = \max(f(x, y), f(x + 1, y), f(x, y + 1), f(x + 1, y + 1)) \quad (3)$$

Where:

- *max_pool(x, y)* represents the output value at axes *(x, y)* in the resulting pooled feature map.

- *f(x, y)* represents the input value at axes *(x, y)* in the input feature map [35].

• **ReLu Activation Function:**

An activation function that may prevent vanishing gradients is the ReLU. This interpretation focuses on the argument's positive axes [18].

$$f(x)_{Relu} = Max(0, x) \quad (4)$$

3.5 Proposed PneuML Sequential CNN Model

The crafted PneuML architecture employed a 12-layer sequential CNN model, as depicted in Fig. 3. The PneuML architecture comprises a lone input layer featuring an image size of 224x224, three convolutional layers (Conv2D) utilizing a Rectified Linear Unit (ReLU) activation function, three maximum pooling layers (Maxpool2D), a flattening layer, two dense layers with ReLu activation, an incorporated dropout layer, and an output layer implementing a SoftMax function.

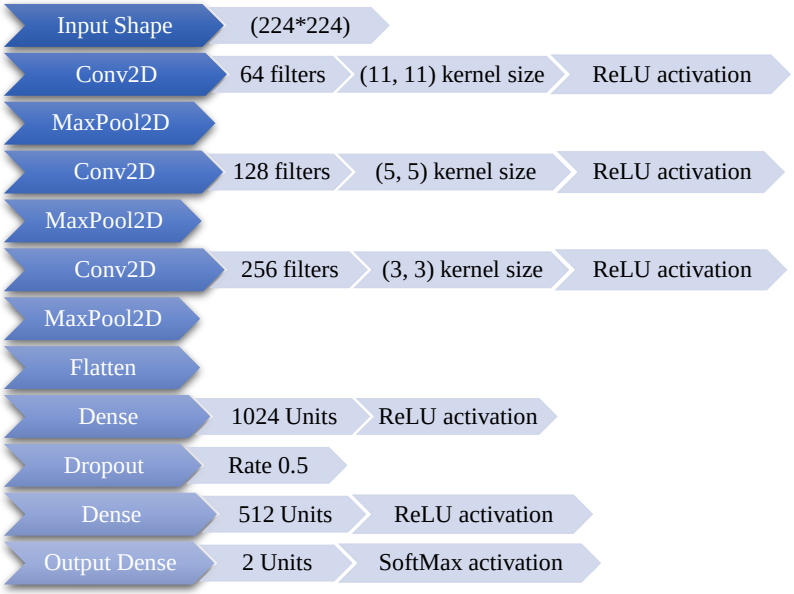

Fig. 3. PneuML Sequential CNN Model Architecture

• **Conv2D:** The convolutional layer is where the model's learning begins, and the convolution process occurs. It includes the kernel and filtering components. In convolutional layers, we filter abstract features and construct a feature map. For the first, second, and third convolutional layers, in our PneuML technique, we used 64 filtrations, 128 filtrations, and 256 filtrations, with kernel sizes of 11x11, 5x5, and 3x3, respectively. Deep learning neural networks frequently employ the ReLu activation function. It is a piecewise linear function that replaces all negative values with 0 while leaving positive values unchanged [34].

• **Maxpool2D:** It is a down-sampling method used to lessen the spatial dimensions of the feature maps obtained in Conv2D. The idea behind maximum pooling is to preserve the most important features while reducing the features and computations. Max pooling takes the maximum value from each miniature, non-overlapping section of the input feature map. The primary use of max pooling is to decrease overfitting and computational complexity. PneuML utilized one Maxpool2D layer for each Conv2D layer in its representation [13, 35].

• **Flatten layer:** The flattened layer converts a multi-dimensional tensor into a single flat vector. A single row preserves the shape of the input.

• **Dense layer:** Dense layers, sometimes called "fully connected layers," are employed in deep learning neural networks. An example of a dense layer is one in which each neuron connects with all of the neurons in the layer below it. This layer performs a dot product between its inputs and a set of weights. PneuML used a highly linked layer with 1024 neuron units for dense layer 1 and 512 neuron units for dense layer 2 [18].

• **Dropout layer:** While each network forward passes, the Dropout layer randomly sets a fraction of the input neurons to zero. It forces the network to rely on multiple neurons to make predictions, reducing the reliance on any particular neuron and preventing overfitting. At each training iteration, the network drops a different set of neurons, resulting in training on various combinations of neurons. Our research uncovered that the dropout rate equals 0.5 [19].

The PneuML training involved using an Adam optimizer with a learning rate of 0.002 and a mean squared logarithmic error loss for 20 epochs, with a batch size of 64. Table 2 outlines the hyper-parameters defined, including the optimizer, the learning rate, the loss function, the batch size, and the epochs.

Table 2.

Hyper-parameters for PneuML and Pre-trained CNNs

Hyper-Parameter	Instance
Optimizer	Adam
Learning Rate	0.002
Loss Function	Mean Squared Logarithmic Error Loss
Batch Size	64
Epochs	20

3.6 Performance Metrics

Metrics employed to assess the effectiveness of an ML or CNN model are known as performance metrics. Classification models primarily use these metrics for evaluation [32]. Examples include accuracy, precision, sensitivity, specificity, recall, and F1 scores. The investigation prioritized the accuracy and F1 scores above the others since they are the most often used and dependable.

4. Results and Discussion

The ML classifiers used for research were SVM, RF, Xgboost, KNN, Gradient Boost, and Adaboost. Fig. 4 displays the accuracy and F1 score achieved by each ML classifier.

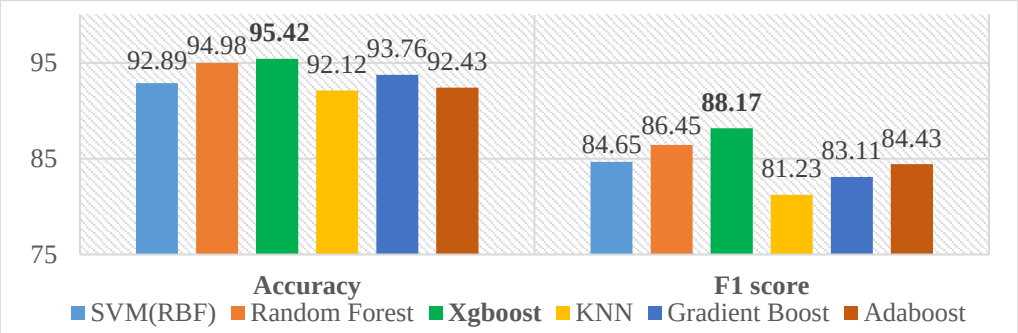

Fig. 4. ML clasifiers outcomes

Using ML algorithms such as SVM, Random Forest, Xgboost, KNN, Gradient Boost, and Adaboost on X-rays for diagnosing pneumonia is a well-established approach in the medical field. The findings suggest that Xgboost demonstrated the best level of accuracy, with a rate of 95.42%. It was closely followed by Random Forest, which achieved an accuracy rate of 94.98%. The XGBoost algorithm exhibited superior performance by attaining the highest F1 score of 88.17%, showcasing its powerful prediction capabilities. Gradient Boost and Adaboost achieved significant results, with F1 scores of 83.11% and 84.43%, respectively. The SVM using the Radial Basis Function (RBF) kernel exhibited a commendable accuracy of 92.89%. However, its F1 score was comparatively lower, measuring at 84.65%. The KNN algorithm showed a modest lag in both accuracy and F1 score, achieving 92.12% and 81.23%, respectively. The results underscore the diverse capabilities of distinct ML algorithms when assessed against the specified measures. This research encompassed a comparative analysis of various classifiers, revealing XgBoost's exceptional performance in pneumonia identification.

These findings underscore the superiority of XgBoost over RF and other ML classifiers. A plausible explanation lies in XgBoost's intricate nature, enabling it to capture subtle dataset patterns. The outcomes suggest XgBoost as the preferred option over Random Forest, potentially attributed to its adeptness in mitigating challenges like noisy data and overfitting common in classification tasks. However, a cautious interpretation of these findings is warranted due to dataset limitations.

The research initially established the suitability of CNNs as the optimal choice for classifying medical images. The investigation delved into the

development of PneuML, a custom CNN architecture, in comparison to established models, including VGG16, VGG19, InceptionV3, DenseNet121, and ResNet50, each configured with the specified hyperparameters. The outcomes, presented in Fig. 5, unveiled compelling insights into the effectiveness of these CNN algorithms in terms of accuracy and F1 score.

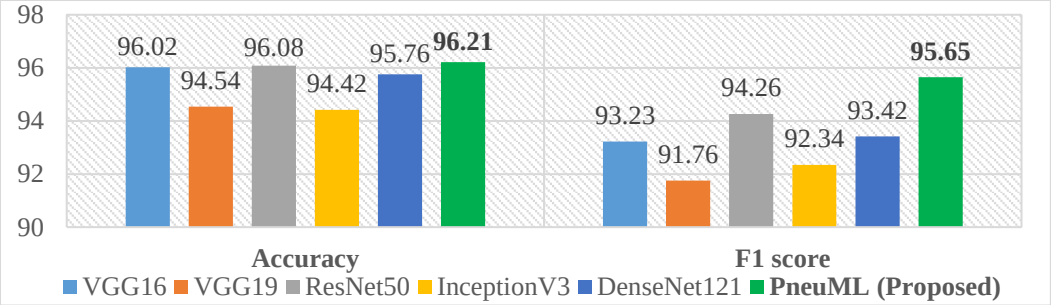

Fig. 5. PneuML Sequential CNN and Pre-trained CNN algorithms outcomes

Among the models under consideration, VGG16 and ResNet50 had notable accuracy rates of 96.02% and 96.08%, respectively. Moreover, as mentioned earlier, the models exhibited noteworthy F1 scores of 93.23% and 94.26%, indicating their efficacy in the assigned undertaking. The InceptionV3 and DenseNet121 models demonstrated strong performance, achieving accuracy scores of 94.42% and 95.76%, respectively. Additionally, these models exhibited equivalent F1 values of 92.34% and 93.42%. Significantly, the PneuML algorithm, as recommended, demonstrated better performance in the given task, surpassing all other algorithms.

PneuML CNN emerged with an impressive accuracy of 96.21% and an F1 score of 95.65%, reflecting its proficiency in accurately identifying instances of pneumonia. Particularly noteworthy, PneuML CNN surpassed ResNet50 with slightly superior accuracy and F1 score, achieving heightened effectiveness and efficiency. ResNet50, on the other hand, exhibited commendable results with a higher accuracy rate and an F1 score. Collectively, the results underscored the exceptional performance of both ResNet50 and PneuML CNN, positioning them as superior choices in accuracy and efficiency compared to other pre-trained CNN models examined. The proposed PneuML sequential CNN was subjected to rigorous validation through a two-class classification task on a generated dataset. Notably, the number of epochs utilized in the PneuML CNN model significantly impacted its accuracy. Incrementing the number of epochs from one to twenty yielded a commensurate increase in accuracy, a trend evident in the accuracy curve shown in Fig. 6. The findings suggest that the model's performance stabilized, demonstrating satisfactory effectiveness. Moreover, the investigation into loss dynamics revealed a reciprocal relationship with the number of epochs. As

portrayed in Fig. 6, the loss magnitude progressively diminished as the epoch count escalated from one to twenty, further underscoring the model's refinement and convergence with extended training.

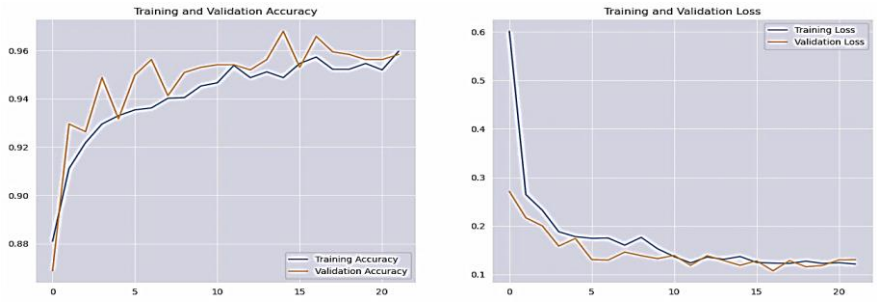

Fig. 6. Accuracy and Loss of the Proposed PneuML Sequetical CNN

The confusion matrix of the PneuML is presented in Fig. 7. The confusion matrix exhibits a robust classification performance, indicating the model's efficacy in accurately detecting both normal and positive class pictures.

TN (1887)	FP(74)
FN(84)	TP(1845)

Fig. 7. Confusion matrix of the Proposed PneuML Sequetical CNN

Table 3 compares the proposed PneuML CNN approach to specific alternatives. Researchers preferred accuracy as their primary performance metric, so the comparison is based on accuracy.

Table 3

Comparison of the Proposed PneuML Sequential CNN

Method	**Accuracy (%)**	**F1 Score (%)**	**References**
SVD-CLAHE Boosting + ResNet-50 + BWCCE	94.00	95.00	[27]
Transfer Learning + Inception V3	92.80	-	[28]
Mobile Freeze-Net	93.00	94.00	[29]
CNN	93.75	94.00	[30]
PneuML - Sequential CNN Architecture (Ours)	96.21	95.65	

The comparison table entailed a comprehensive assessment of various techniques for categorizing X-ray images sourced from the LDOCTCXR and Balanced Augmented COVID-CXR datasets, primarily focusing on pneumonia diagnosis. The investigation highlighted the importance of approaches such as

SVD-CLAHE, ResNet-50, and BWCCE in producing robust findings. The effectiveness of the Inception V3 transfer learning technique was notable, demonstrating the need to leverage pre-trained models for accurate predictions. The Mobile Freeze-Net method showed the efficacy of mobile-focused architectures. The use of CNN improved the dependability of proven deep learning algorithms. The PneuML sequential CNN Architecture approach achieved 96.21% accuracy and a 95.11% F1 score. These outcomes underscore the potential advantages of employing tailored architectures designed to analyze pneumonia X-ray images. Overall, the investigation showcased the impact of various methodologies and highlighted the diverse range of approaches adopted by researchers.

This research uses ML techniques based on X-ray images to diagnose pneumonia, including pre-trained CNNs and a novel PneuML sequential model. The study places particular emphasis on doing a comparative analysis of different classifiers. The main aim of the research was to ascertain the strategy that had the most efficacy in detecting pneumonia. The investigated ML methods included SVM, Random Forest, XgBoost, KNN, Gradient Boost, and Adaboost. This research revealed that the XgBoost algorithm exhibited superior performance to other algorithms, resulting in a notable increase in classification accuracy.

In contrast, the Random Forest model ranked second with a somewhat reduced efficacy. These findings highlight XgBoost's dominance over RF and other ML classifiers. XgBoost's excellent success might be due to its intricate approach, which enables it to identify small patterns within the dataset efficiently. This capacity is advantageous in dealing with issues like noisy data and overfitting, typical in classification work. However, owing to possible limitations in the dataset, it is critical to take care when interpreting these findings.

The research subsequently turned its attention to the usefulness of CNNs for classifying generated dataset images. Pre-trained CNNs such as VGG16, VGG19, InceptionV3, DenseNet121, and ResNet50 were employed, each with its own set of hyperparameters. ResNet50 performed well in comparison to the other CNNs. The findings revealed that the ResNet50 performed excellently over other CNN models, establishing him as the better alternative.

The results revealed that PneuML CNN achieved outstanding effectiveness, showcasing its proficiency in accurately identifying instances of pneumonia. Notably, PneuML CNN outperformed ResNet50 and XgBoost, indicating heightened efficacy and efficiency. The generated dataset validated the PneuML sequential CNN in a two-class classification test. Significantly, the number of epochs in the PneuML CNN model affected its accuracy. The increase in epochs from one to twenty resulted in a corresponding increase in accuracy. PneuML shows that extensive training stabilized model performance, proving its efficacy. Loss dynamics also showed a reciprocal correlation with epochs. As the number of

epochs increased, the magnitude of loss decreased, highlighting the model's improvement and convergent convergence during training.

5. Conclusion

Pneumonia is still a significant global health issue, increasing mortality rates. This research aims to address it through the development of ML and refined PneuML image classification methodology. The study endeavored to develop an image classification model that predicted X-ray images within two primary classes, pneumonia and normal, achieving its intended goal. The outcomes obtained from the developed PneuML model were promising, displaying the robust effectiveness of the PneuML CNN, which reached an accuracy rate of 96.21% and an impressive F1 score of 95.65% and presented its potential as a valuable asset. The PneuML sequential CNN performed better than both pre-trained CNNs, with ResNet50 being best served, and ML, where XgBoost ML classifiers were best suited, providing a comparative analysis of CNNs and ML. The experiment involved the amalgamation of two publically accessible datasets to generate a unique pneumonia and normal images dataset. While successful for the task at hand, sequential CNNs have limits in acquiring entire contextual information, identifying complicated patterns, handling long-range associations, and interpreting. Furthermore, their effectiveness is often dependent on precise hyper-parameter adjustment. Integrating cutting-edge techniques, including concepts such as class weights, holds promise for further performance enhancements. Despite the favorable findings, future research on a vast dataset remains essential to establish a more robust assessment of performance metrics.

REFERENCES

[1]. *The top 10 causes of death*, "The top 10 causes of death," Dec. 09, 2020. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (accessed on 11 March 2023)

[2]. *The Top 10 Deadliest Diseases in the World*, "The Top 10 Deadliest Diseases in the World." https://www.healthline.com/health/top-10-deadliest-diseases (accessed on 12 March 2023)

[3]. *G. Mackenzie*, "The definition and classification of pneumonia," in Pneumonia, **vol. 8**, no. 1, Aug. 2016. doi: 10.1186/s41479-016-0012-z

[4]. *D. S. Kermany, K. Zhang, M. Goldbaum*, "Large Dataset of Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images", Mendeley Data, V3, 2018. doi: 10.17632/rscbjbr9sj.3

[5]. *M. E. H. Chowdhury et al.*, "Can AI Help in Screening Viral and COVID-19 Pneumonia?" in IEEE Access, **vol. 8**, 2020, pp. 132665–132676. doi: 10.1109/access.2020.3010287

[6]. *H. Verma et al.*, "Local Binary Patterns Based on Neighbor-Center Difference Image for Color Texture Classification with Machine Learning Techniques," in Wireless Communications and Mobile Computing, **vol. 2022**, Aug. 2022, pp. 1–11. doi: 10.1155/2022/1191492

[7]. *C.E. da S. Santos, L. dos S. Coelho, and C.H. Llanos*, "Nature inspired optimization tools for SVMs - NIOTS," in MethodsX, Vol. 8, 2021, pp. 101574. https://doi.org/10.1016/j.mex.2021.101574

[8]. *H. Arslan and H. Arslan,* "A new COVID-19 detection method from human genome sequences using CpG island features and KNN classifier," in Engineering Science and Technology, **vol. 24, no. 4,** Aug. 2021, pp. 839–847. doi: 10.1016/j.jestch.2020.12.026

[9]. *K. Shaheed, P. Szczuko, Q. Abbas, A. Hussain, M. Albathan*, "Computer-Aided Diagnosis of COVID-19 from Chest X-ray Images Using Hybrid-Features and Random Forest Classifier," in Healthcare, **Vol. 11**, 2023, pp.837. https://doi.org/10.3390/healthcare11060837.

[10]. *H. V. Donga, J. S. A. N. Karlapati, H. S. S. Desineedi, P. Periasamy, and S. TR,* "Effective Framework for Pulmonary Nodule Classification from CT Images Using the Modified Gradient Boosting Method," in Applied Sciences, **vol. 12, no. 16**, Aug. 2022, pp. 8264. doi: 10.3390/app12168264

[11]. *H. Nasiri and S. Hasani,* "Automated detection of COVID-19 cases from chest X-ray images using deep neural network and XGBoost," in Radiography, **vol. 28, no. 3**, Aug. 2022, pp. 732–738. doi: 10.1016/j.radi.2022.03.011

[12]. *E. Sevinç*, "An empowered AdaBoost algorithm implementation: A COVID-19 dataset study," in Computers & Industrial Engineering, **vol. 165**, Mar. 2022, pp. 107912. doi: 10.1016/j.cie.2021.107912

[13]. *D. Alapat, M. Menon, S. Ashok,* "A Review on Detection of Pneumonia in Chest X-ray Images Using Neural Networks," in Journal of Biomedical Physics and Engineering, **vol. 12, no. 6,** Dec. 2022. doi: 10.31661/jbpe.v0i0.2202-1461

[14]. *K. Simonyan, and A. Zisserman*, "Very deep convolutional networks for large-scale image recognition," arXiv, 2014, arXiv:1409.1556.

[15]. *Md. B. Hossain, S. M. H. S. Iqbal, Md. M. Islam, Md. N. Akhtar, and I. H. Sarker,* "Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images," in Informatics in Medicine Unlocked, **vol. 30**, 2022, pp. 100916, doi: 10.1016/j.imu.2022.100916

[16]. *D. Miserlis et al*., "Benchmarking EfficientNetB7, InceptionResNetV2, InceptionV3, and Xception Artificial Neural Networks Applications for Aortic Pathologies Analysis," in Journal of Vascular Surgery, **vol. 77, no. 6,** Jun. 2023, pp. e345, doi: 10.1016/j.jvs.2023.03.475

[17]. *L. Kong, and J. Cheng,* "Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion," in Biomedical Signal Processing and Control. **77**, 2022, pp. 103772. https://doi.org/10.1016/j.bspc.2022.103772

[18]. *L. Alzubaidi et al.,* "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," in Journal of Big Data, **vol. 8, no. 1,** Mar. 2021. doi: 10.1186/s40537-021-00444-8

[19]. *P. Szepesi and L. Szilágyi,* "Detection of pneumonia using convolutional neural networks and deep learning," in Biocybernetics and Biomedical Engineering, **vol. 42**, **no. 3,** Jul. 2022, pp. 1012–1022. doi: 10.1016/j.bbe.2022.08.001

[20]. *D. Avola, A. Bacciu, L. Cinque, A. Fagioli, M. R. Marini, and R. Taiello*, "Study on transfer learning capabilities for pneumonia classification in chest-x-rays images," in Computer Methods and Programs in Biomedicine, **vol. 221**, Jun. 2022, pp. 106833. doi: 10.1016/j.cmpb.2022.106833

[21]. *J. Liu, J. Qi, W. Chen, and Y. Nian*, "Multi-branch fusion auxiliary learning for the detection of pneumonia from chest X-ray images," in Computers in Biology and Medicine, **vol. 147**, Aug. 2022, pp. 105732. doi: 10.1016/j.compbiomed.2022.105732

[22]. *G. Srivastava, N. Pradhan, and Y. Saini*, "Ensemble of Deep Neural Networks based on Condorcet's Jury Theorem for screening Covid-19 and Pneumonia from radiograph images,"

in Computers in Biology and Medicine, **vol. 149**, Oct. 2022, pp. 105979. doi: 10.1016/j.compbiomed.2022.105979

[23]. *Y. Qu, Y. Meng, H. Fan, and R. X. Xu*, "Low-cost thermal imaging with machine learning for non-invasive diagnosis and therapeutic monitoring of pneumonia," in Infrared Physics & Technology, **vol. 123**, Jun. 2022, pp. 104201. doi: 10.1016/j.infrared.2022.104201

[24]. *A. K. Singh, A. Kumar, M. Mahmud, M. S. Kaiser, and A. Kishore*, "COVID-19 Infection Detection from Chest X-Ray Images Using Hybrid Social Group Optimization and Support Vector Classifier," in Cognitive Computation, Mar. 2021. doi: 10.1007/s12559-021-09848-3

[25]. *S. Kumar, H. Kumar*, "Lungcov: A diagnostic framework using machine learning and Imaging Modality," in International Journal on Technical and Physical Problems of Engineering (IJTPE), **Issue 51, Volume 14**, **Number 2,** June 2022. pp. 190-199. http://mail.iotpe.com/IJTPE/IJTPE-2022/IJTPE-Issue51-Vol14-No2-Jun2022/23-IJTPE-Issue51-Vol14-No2-Jun2022-pp190-199.pdf

[26]. *P. K. Wong et al.,* "Automatic detection of multiple types of pneumonia: Open dataset and a multi-scale attention network," in Biomedical Signal Processing and Control, **vol. 73**, Mar. 2022, pp. 103415. doi: 10.1016/j.bspc.2021.103415

[27]. *S. Roy, M. Tyagi, V. Bansal, and V. Jain*, "SVD-CLAHE boosting and balanced loss function for Covid-19 detection from an imbalanced Chest X-Ray dataset," in Computers in Biology and Medicine, **vol. 150**, Nov. 2022, pp. 106092. doi: 10.1016/j.compbiomed.2022.106092

[28]. *D. S. Kermany et al*., "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," Cell, **vol. 172, no. 5,** Feb. 2018, pp. 1122-1131.e9. doi: 10.1016/j.cell.2018.02.010

[29]. *S. Roy and R. Khurana*, "Mobile Freeze-Net with Attention-based Loss Function for Covid-19 Detection from an Imbalanced CXR Dataset," Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing, Mar. 2023. doi: 10.1145/3555776.3577825

[30]. *S. Singh and B. K. Tripathi*, "Pneumonia classification using quaternion deep learning," Multimedia Tools and Applications, **vol. 81, no. 2,** Oct. 2021, pp. 1743–1764. doi: 10.1007/s11042-021-11409-7

[31]. "Balanced Augmented Covid CXR Dataset," Balanced Augmented Covid CXR Dataset | Kaggle. /datasets/tr1gg3rtrash/balanced-augmented-covid-cxr-dataset

[32]. *A. P. Ayshath Thabsheera, T. M. Thasleema, and R. Rajesh*, "Lung Cancer Detection Using CT Scan Images: A Review on Various Image Processing Techniques," Data Analytics and Learning, Nov. 2018, pp. 413–419. doi: 10.1007/978-981-13-2514-4_34

[33]. *I. H. Sarker*, "Machine Learning: Algorithms, Real-World Applications and Research Directions," SN Computer Science, **vol. 2**, no. 3, Mar. 2021. doi: 10.1007/s42979-021-00592-x

[34]. *C. S. Goodrich*, "Convolution equations with variable time nonlocal coefficients," Applied Mathematics Letters, **vol. 145**, Nov. 2023, pp. 108756. doi: 10.1016/j.aml.2023.108756

[35]. *A. Zafar et al*., "A Comparison of Pooling Methods for Convolutional Neural Networks," Applied Sciences, **vol. 12, no. 17,** Aug. 2022, pp. 8643. doi: 10.3390/app12178643