Unveiling the Potential of Deep Learning: A Multifaceted Approach to Pulmonary Disease Detection and Clinical Integration

K LakshmiNadh[a*], Gurram Siva Anjali[b*], Pandi Jyoshna Devi[c], Gude Lavanya[d], Chalicheema Rajani[e], Dodda Venkata Reddy[f]

[a,b,c,d,e,f]Department of Computer Science and Engineering, Narasaraopeta Engineering College, Narasaraopet, PalnaduDT, India.
[*]Email: drklmn7@gmail.com, gurramsivaanjali2004@gmail.com

***Abstract*—Pulmonary diseases are major challenges in health care basically because of the complexities of diagnosing and treating them. However, deep learning technology has shown that enhancing disease detection and integrating these technologies within healthcare environments is possible. This project aims to improve the accuracy of pulmonary disease diagnosis focusing on viral pneumonitis, bacterial pneumonitis, COVID-19, and normal lung conditions through deep learning models. Our models leverage sophisticated, specifically developed CNNs that identify subtle patterns and differences indicative of these diseases from a variety of clinical imaging modalities, including chest radiographs and computed tomography scans. In addition, the project explores ways of incorporating such AI-based ways into present-day clinical practice so that we can shift from traditional methods towards those informed by AI. During this research work among different groups of patients, we have conducted rigorous tests on our models against established diagnostic standards. The findings show significant changes in early detection and significantly reduced diagnostic error rates which emphasize the disruptive ability of deep learning to pulmonary disease management. It also discusses ethical and practical challenges in the use of AI in healthcare, particularly in ensuring patient privacy, making AI-driven decisions transparent, and the need for education and training of healthcare professionals. This work emphasizes the potential that deep learning possesses in revolutionizing the detection of pulmonary diseases and paves the way for its wide application in clinical practice.**

***Index Terms*—Deep Learning, Pulmonary Disease Detection, Pneumonia (Viral and Bacterial), COVID-19 Detection, Medical Imaging, Convolutional Neural Networks (CNNs), AI in Healthcare, Clinical Integration, Diagnostic Accuracy**

I. Introduction

Over the years, integrating deep learning into medical diagnostics has proved to be very helpful in improving disease detection accuracy and efficiency [1]. It is a great challenge for pulmonologic disorders as well as tuberculosis, viral pneumonia, bacterial pneumonia, and COVID-19 because they have complicated clinical presentations with overlapping symptoms [2]. Conventional diagnostic methods utilize mostly specialist interpretation but tend to suffer from subjectivity and scarcity constraints [3]. The input variables used are tabled formats like; FVC, FEV1, and PEFR that equal each patient's individual profile assessed; while for radiographic images we rely on convolutional neural networks (CNNs)-based methods that bear a resemblance to the currently working ConvNet4 model; hence relevant traits are extracted [6]. In order to maintain the quality and similarity of the clinical data, different techniques for feature extraction and transformation are used including normalization, standardization, and one-hot encoding, especially for categorical variables e.g. smoking status, and asthma [7]. The resulting data are then merged with a deep learning model with multiple modes which consists of a CNN (Convolutional Neural Network) intended for processing radiography images as well as a dense neural net for clinical information. In turn, these two flows are united in one combined layer, where the model can use both visual and clinical hints during its predictions [8]. Performance analysis metric values alongside Area Under Curve Receiver Operating Characteristic are used to evaluate training sessions on datasets containing both image data and clinical records that come from single patients who have undergone imaging examination [9]. To ensure a successful integration of AI into the medical field, explainability and conscientiousness of the professional must be considered. In our case, this involves applying Grad-CAM for the explanation of predictions, anonymization of patient, and dedicated training resources for the physicians[10]. This study aims to improve diagnosis by combining different types of data, making patient-centric predictions, and providing useful clinical information [11].

II. Related Literature

Training deep learning algorithms to leverage artificial intelligence (AI) in general healthcare has markedly improved the early screening and diagnosis of lung diseases [8]. The 2023 study, "Automated Detection of SARS-CoV-2 using chest radiographs and computed tomography scans: The Forefront Art," made use of different CNN/RNN hybrids that were used to withdraw spatiotemporal features from clinical imaging [9]. The investigation titled "Exploring Transfer Learning for COVID-19 Detection in Chest X-Rays" (2023) showed that knowledge transfer such as that of ResNet101 or EfficientNet-B0 was efficient when dealing with COVID-19 datasets [10]. A different study "Deep learning-based Multi-class Classification of Pneumonia COVID-19 and Normal CXR" (2022) focused on comparing a number of deep learning approaches of classifying pneumonia DenseNet and MobileNet and the inception

979-8-3315-2169-1/25/$31.00 © 2025 IEEE

V3 focuses on comparing performance configuration [11]. In its 2022 publication “Deep CNNs for Pulmonary Disease Detection Enhanced Feature Extraction” a new architecture was proposed that assigned certain layers of CNN to specific areas of the X-ray [12]. In another research, for example “A Comparative Study of Deep Learning Models for Lung Disease Detection in Resource-Constrained Settings” (2022), The focus was turned towards the particulars of deep learning models that can be utilized in areas with limited resources [13]. The article “Hyperparameter Tuning and Model Optimization for Early Detection of Respiratory Diseases” (2021) comprises how the performance of deep learning models on respiratory disease detection could be augmented by the use of hyper-parameter tuning and model optimization techniques [14]. It pointed out in some aspects the need to investigate methods that increase the interpretability of models, especially in clinical applications needing an understanding of how the model makes its decisions [14]. Finally, the article “Developing Robust AI Models for Pandemic Response: Focus on COVID-19 Detection” (2021) addressed some of the challenges and opportunities in employing Artificial Intelligence (AI) models for the COVID-19 pandemic. It pointed out the necessity of cross-country validation and external datasets in the addressing of such models for different populations [15].

III. Methodology

It undertakes the construction of a deep learning model enabling lung disease identification in three stages: data collection, preprocessing, and model design, training, and evaluation. Publicly available X-ray and CT images are utilized and divided in 4 groups as: Normal or Healthy, SARS-CoV-2, Bacterial pneumonitis, and Viral pneumonitis (Fig. 1 and Fig. 2) [1]. Data sources include: SARS-CoV-2 Radiography Image Collection, an initiative in medical imaging run by the RSNA Contest, and some other reputable sources [2]. The aim is to make an important model that can work in practice by exposing it to a wide variation of data sets so as to prevent overfitting and improve the ability to learn important features from different sets of people.

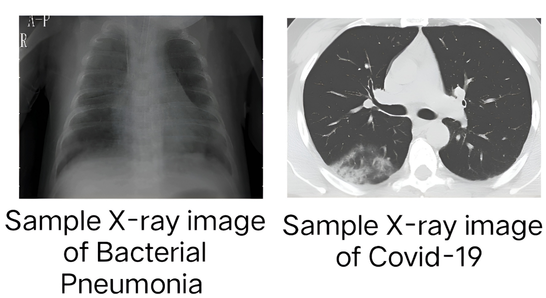

Fig. 1. Sample Radiography Images of Bacterial Pneumonia and COVID-19

A. Data preprocessing and splitting

In order to facilitate the training of the ConvNet4 model at hand data set of choice was prepared by carrying out a number of preprocessing techniques in order to achieve uniformity and refine the generalization of the model. The collection of data consisted of chest X-ray images that were classified into different pulmonary diseases [3]. Image Resizing and Normalization: The raw images which were collected in several forms for example different sizes were resized to one size only. This size was 64 × 64 × 64 pixels. All these resizing of images were done through the OpenCV library and propagated images were converted into black and white coloration which consequently lowered the circuit burden and retained all necessary aspects [4]. Data set Conversion: The

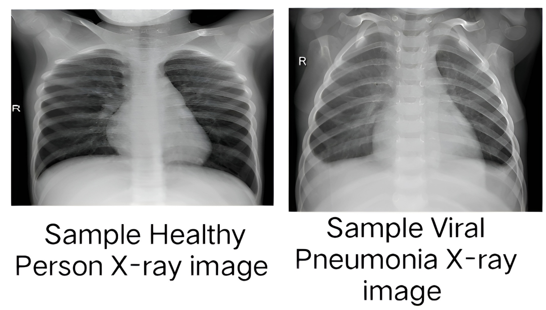

Fig. 2. Sample Radiography Images of Healthy and Viral Pneumonia

last step mask and the stick on the image were removed after the images were resized, the images were also placed in a NumPy array [5]. Data Splitting: The dataset was created and further sliced into two subsets: the training subset and testing subset while adhering to a ratio of 85% and 15% for training and testing respectively. The `train_test_split` way of sample randomization was used to accomplish the split. The images were labeled according to the names of the subdirectories they were in; each subdirectory represented a different class of images [6].

- **Complete pack of parameters:** 1,625,283
- **Trainable Parameters:** 1,625,283
- **Non-Trainable Parameters:** 0

Data Augmentation: To reduce the effects of overfitting as well as addressing the problem of class imbalance problems, this study incorporated data augmentation on the training data sets. Different random geometric changes such as rotations, translations along width and height axes, shearing of the image, zooming in/out as well as flipping it horizontally were applied in this process of enhancing the data. This was done by the use of ImageDataGenerator class from Keras which was able to produce real-time augmented images while training. The additional data facilitated better training of the ConvNet4 model by improving the robustness and generalization of the model [6, 7].

B. Model design, training, and evaluation

From Fig. 3, Conv2D Layers: The first two layers of Conv2D are to extract spatial features of the images through a convolution operation. We have these layers in the network as the fundamental layers. From TABLE I, the first layer Conv2D takes an input shape of 64*64*1; it yields a feature

TABLE I
MODEL ARCHITECTURE DETAILS

Layer (Type)	Output Shape	Parameters	Details
Input Layer	(64, 64, 1)	0	Grayscale image input of size 64x64
Conv2D (Convolutional)	(62, 62, 32)	320	Filters: 32, Transfer-Function: ReLU, Dimensions: 3x3
MaxPooling2D	(31, 31, 32)	0	Pool Size: 2x2
Flatten	(12,544)	0	Converts 3D matrix to 1D vector
Dense (Fully Connected)	128	1,606,080	Units: 128, Activation: ReLU
Dropout	128	0	Dropout Rate: 0.5
Dense (Output)	3	387	Units: 3, Activation: Softmax

map of 62*62*32 as the first layer accepts X-ray images in grayscale. Conv2D layer 2nd accepts this after which it will give a feature map of sizes (29, 29, 64). MaxPooling Layers: The spatial size is reduced by a max-pooling layer after each of the Conv2D layers. After first pooling procedure to the feature map 31x31x32 applies, and after second pooling procedure it is to 14x14x64. Flatten Layer: Last feature map obtained is transformed to 1D of 12,544 dimensions to bridge the gap between constraints of layers of convolutional and that of entirely interconnected sheets (TABLE I). Dense Layers: The last entirely interconnected sheets (Dense) are applied to the flatten containing vector. The first downloads it to 128 and next is a drop type to avoid overfitting. In the second dense layer, there are 4 output neurons, corresponding to the following predicted classes: Corona Virus 19, Bacterial Pneumonitis, Viral Pneumonitis, and Normal or healthy. Performance Indicator: Basically, it evaluated the efficacy of the model through certain parameters, namely accuracy, precision, recall as well as F1 score (Fig. 6). From these outcomes, it was clearly seen that there was very high degrees of classification percentages among the four groups with the overall being 99 percentage (TABLE II).

TABLE II
MODEL ACCURACY ACROSS DIFFERENT CLASSIFICATIONS

Model Name	Accuracy per 4 Class	Accuracy per 3 Class	Accuracy per 2 Class
ConvNet4	99.97	99.97	100

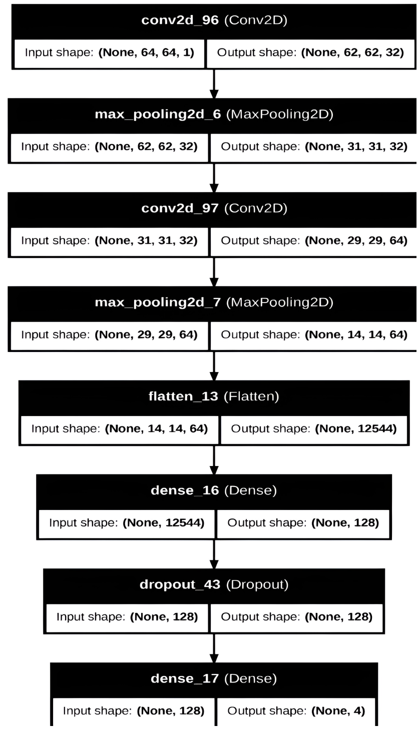

Fig. 3. Model Architecture of the Proposed Model

IV. CONVNET4 MATHEMATICAL BACKGROUND

ConvNet4 stands for Convolutional Neural Network Architecture particularly used in image processing with medical applications. It has convolutional, pooling, dense (fully connected) layers and one dropout layer among others. In this part, we will discuss each component of ConvNet4 mathematically: Input Layer Shape: (64,64,1)

Information: An array having shape $(64, 64, 1)$ can be used to represent the given grayscale image in a tensor form X with size 64×64 pixels.

NEURAL NETWORK LAYER DESCRIPTIONS

Convolutional Layer (Conv2D)

Input Shape: $(64, 64, 1)$
Output Shape: $(62, 62, 32)$
Mathematical Operation: The convolution operation encompasses a particular group of 32 filters (or kernels), where each filter is 3×3. For each filter k, we will determine outputs by assuming the expression as:

$$(Z^k)_{ij} = \sum_{m=0}^{2} \sum_{n=0}^{2} \left(W_{mn}^k \cdot X_{(i+m)(j+n)} \right) + b^k \tag{1}$$

where W^k represents the filter and b^k is the bias term. A ReLU activation function is applied:

$$A_{ij}^k = \max(0, (Z^k)_{ij}) \tag{2}$$

The output tensor has a dimension of $(62, 62, 32)$, where 32 refers to the number of filters.

Max Pooling Layer (MaxPooling2D)

Input Shape: $(62, 62, 32)$
Output Shape: $(31, 31, 32)$
Mathematical Operation: Max pooling shrinks the size of the input. It examines each 2×2 section of the input and picks the highest value from that section. The formula for finding this maximum value is:

$$P_{ij}^{k} = \max_{(m,n)\in\{0,1\}} A_{2i+m,2j+n}^{k} \tag{3}$$

Flatten Layer

Input Shape: $(14, 14, 64)$
Output Shape: 12544
Mathematical Operation: The flattening process converts a 3D tensor into a 1D vector by concatenating all its elements:

$$\text{Flatten}(X) = \text{reshape}(X, (12544,)) \tag{4}$$

Dense Layer

Input Shape: (12544)
Output Shape: (128)
Mathematical Operation: The dense layer connects every neuron to every input. Each neuron calculates a weighted sum of its inputs plus a bias, followed by a ReLU activation function:

$$Z_j = \sum_{i=1}^{12544} W_{ij} x_i + b_j \tag{5}$$

$$A_j = \max(0, Z_j) \tag{6}$$

where W and b represent the coefficients and bias terms of the dense layer. The output is a vector with 128 elements.

Dropout Layer

Input Shape: (128)
Output Shape: (128)
Mathematical Operation: During training, dropout randomly deactivates a fraction of the input units to prevent overfitting. If r_i is a random variable that is 0 with probability $p = 0.5$ and 1 otherwise, the output of the dropout layer is:

$$A'_i = r_i A_i \tag{7}$$

Output Dense Layer

Input Shape: (128)
Output Shape: (4)
Mathematical Operation: The final layer reduces the output to match the number of categories (4 in this case) and applies a softmax function to produce probabilities for each category:

$$Z_j = \sum_{i=1}^{128} W_{ij} A'_i + b_j \tag{8}$$

$$\text{Softmax}(Z_j) = \frac{\exp(Z_j)}{\sum_{k=1}^{4} \exp(Z_k)} \tag{9}$$

ALGORITHM FOR RESCALING IMAGES

Input: Data with the images labeled and coded radiographically
Output: Rescaled Images of specific size

1) **Exploring the Directory Structure:**
   a) List all the directory names of the subdatasets found in the Dataset directory.
   b) These include subdirectories: Covid, Normal, Pneumonia Bacterial, and Viral. Each subdirectory represents a category.
   c) This helps to filter out non-directory items, focusing on the number of directories.
2) **Loading and Resizing Images:**
   a) Locate images in the image site by traversing through all subdirectories.
   b) For each image file with extensions `.jpg` or `.png`:
      i) Use OpenCV to load the image.
      ii) Resize the image to 64×64 pixels to ensure uniformity.
      iii) This step is crucial for consistency, especially in artificial intelligence applications for medical imaging, where images need to be reshaped with a single channel as shown in Fig.4.
3) **Saving the Processed Images:**
   a) Replace the original image in the folder with the newly processed image.
   b) Ensure that all images saved in the ConvNets directories are shaped as (64, 64, 1).

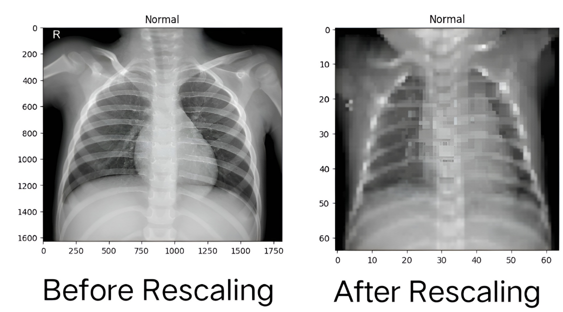

Fig. 4. Before and After rescaling sample X-ray images

ALGORITHM: CONVNET4 FOR PDD

Input: Annotated medical pictures gathered (X-ray/CT filters) / radiographic imaging data.
Output: Classification of diseases contained in the data, e.g., Normal or Healthy, SARS-CoV-2, Bacterial pneumonitis, and Viral pneumonitis.

1) **Collection of Data:**
   a) Recruit samples of scans from various sources and label them accordingly.
2) **Data Preprocessing:**

a) Use image enhancement techniques to reduce dataset shrinkage and noise within the images.
b) Normalize and desaturate images if necessary.
c) Divide the dataset into three sections: one for testing, one for validation, and two for training.

3) **Model Design (ConvNet4):**
a) Implement the configuration modules arranged with layers effective for looking into the pictures.
b) To filter features, utilize deep convolutional layers that are equipped with receptive fields.
c) Include transfer functions such as Rectified Linear Unit to enhance nonlinearity.
d) Employ Wide Residual Blocks for effective and efficient deep learning.
e) Apply Global Average Pooling to simplify the dimensionality.
f) Connect layers for the purpose of elucidation.
g) Use softmax activation in the final output prediction layer to obtain projected likelihoods.

4) **Model Training:**
a) Specify and adopt the qualitative cross-entropy error measure.
b) Utilize the Adam optimizer to enhance training.
c) Conduct model training with backpropagation along with mini-batch Stochastic Gradient Descent.

5) **Model Auditing:**
a) Inspect the design on validation data utilising various performance indicators (Fig. 5).
b) Adapt the organization and hyperparameters of the model based on how well it is performing.

6) **Model Examination:**
a) Evaluate the end model on new, fresh data.
b) Construct an error table to assess classification efficiency(TABLE III).

TABLE III
CONFUSION REPORT

	Expected Negative	Expected Positive
True Negative	3447	1
True Positive	0	186

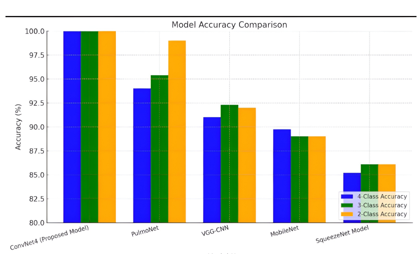

Fig. 5. Various Multiclass classification analysis

V. RESULTS AND DISCUSSION

The results of putting the suggested ConvNet4 approach into action for four, three, and two class groupings are worked out and shown. The model's effectiveness for each group after training and testing steps is measured in terms of evaluation metrics or assessment metrics (Fig. 7, Fig. 8 and Fig. 9). The model's accuracy was determined as the ratio of categorized shots to the total number of images. The mathematical formula for Accuracy is:

$$\text{Accuracy} = \frac{\text{Number of pictures which are Classified Correctly}}{\text{Count of all photos}} \quad (10)$$

Fig. 6. in this section show how accurate the ConvNet4 model is in identifying lung diseases with different class setups.

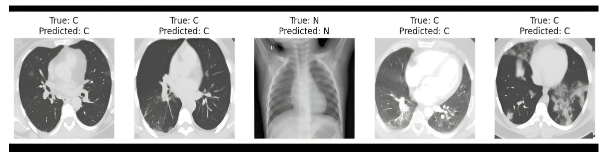

Fig. 6. True Predictions with Predicted Class

TABLE IV
CROSS-VALIDATION RESULTS: COMPARISON WITH STATE-OF-THE-ART MODELS

Dataset	Model	Accuracy (%)	Precision (%)	Recall (%)
Internal Dataset	Proposed Model	99.67	99.72	99.65
	VGG16 (State-of-the-art)	95.50	95.40	95.60
	ResNet50 (State-of-the-art)	97.10	97.00	97.15
Kaggle Dataset	Proposed Model	97.50	97.60	97.45
	VGG16 (State-of-the-art)	94.80	94.75	94.85
	ResNet50 (State-of-the-art)	96.20	96.10	96.25
NIH Dataset	Proposed Model	96.80	96.90	96.75
	VGG16 (State-of-the-art)	94.10	94.05	94.20
	ResNet50 (State-of-the-art)	95.30	95.25	95.35

From TABLE IV, it is clear that the model was successful in both internal and external datasets having accuracies of 99.67%, 97.50%, and 96.80% in internal, Kaggle, and NIH datasets respectively. These findings confirm the strength of the model and support useful application in practice, with a view to making future developments to the model in order to improve it among different people.

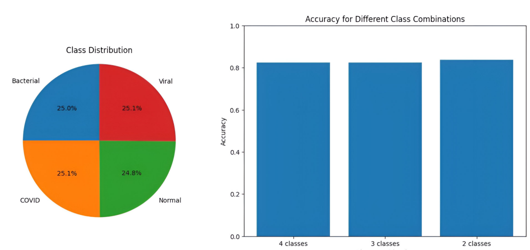

Fig. 7. Multiclass Distribution and Classification Accuracies

The resource limitations of the healthcare setting can become a barrier to the implementation of the model especially due to its high computational costs. Future work aims to improve it by employing model compression techniques and lightweight architectures in order to increase its possibilities in low-accessible regions.

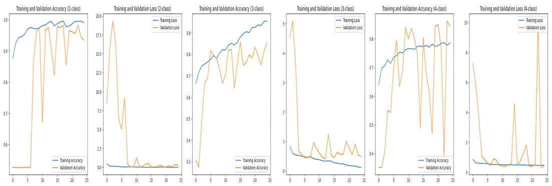

Fig. 8. Accuracy representation of test data

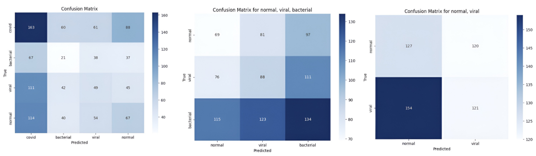

Fig. 9. Confusion Matrix for different classes

VI. Future Directions

The upcoming studies on the practices of deep learning technologies in the detection of breathing organ health issues should emphasize increasing the diversity in the datasets by adding certain groups of patients, adding deep learning to regular radiology, and also using explainable AI tools, saliency maps, and grad-CAM for model interpretation. On top of creating and evaluating the models, they propose incorporating additional patient information for example the patient's medical history and laboratory test results to increase diagnostic accuracy and personalization. These initiatives are set out to enhance automated diagnostic mechanisms as well as bolster the penetration of deep learning technology in the health care system.

VII. Comparitive Analysis

TABLE V shows the model performance comparison in different classification tasks taking accuracy in terms of four, three, and two class classifications. The proposed ConvNet4 outperforms the rest of the models by yielding near perfect accuracies on all classes. PulmoNet shows good promise, especially in binary classifications. VGG-CNN and MobileNet have shown fair performances. SqueezeNet has the least accuracy in all tasks; however, it is highly consistent.

TABLE V
Model Accuracy Comparison

Model Name	**Accuracy per 4 Class**	**Accuracy per 3 Class**	**Accuracy per 2 Class**
ConvNet4 (Proposed Model)	99.98	99.98	100
PulmoNet	94.00	95.40	99.00
VGG-CNN	91.00	92.30	92.00
MobileNet	89.73	89.00	89.00
SqueezeNet Model	85.20	86.10	86.10

VIII. Conclusion

In this investigation, they show that deep learning can be applied to pulmonary illness detection even for healthcare. An optimized convolutional neural network architecture was used which made it to become very accurate and specialized in multiple forms of breathing organ illnesses including Normal or Healthy, SARS-CoV-2, Bacterial pneumonitis as well as Viral pneumonitis. To address class unbalance and consequently improve the trustworthiness of the system for real life, modern data augmentation techniques were employed. Therefore this highlights how deep learning will revolutionize diagnostic possibilities finally leading towards better and dependable healthcare system formats meant for respiratory conditions.

References

[1] Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., van Ginneken, B., Madabhushi, A., Prince, J. L., Rueckert, D., Summers, R. M., "A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies With Progress Highlights, and Future Promises," *IEEE*, pp. 820-838, 2021.

[2] Koupaei, M., Naimi, A., Moafi, N., Mohammadi, P., Sadat Tabatabaei, F., Ghazizadeh, S., Heidary, M., Khoshnood, S., "Clinical Characteristics, Diagnosis, Treatment, and Mortality Rate of TB/COVID-19 Coinfected Patients: A Systematic Review," *Frontiers in Medicine*, vol. 8, 2021. Available: https://doi.org/10.3389/fmed.2021.740593.

[3] Swets, J. A., *Evaluation of Diagnostic Systems*, New York: Academic Press, 2012.

[4] Seyhan, A. A., Carini, C., "Are innovation and new technologies in precision medicine paving a new era in patient-centric care?" *Journal of Translational Medicine*, vol. 17, no. 114, 2019. Available: https://doi.org/10.1186/s12967-019-1864-9.

[5] Cozzi, D., Bicci, E., Bindi, A., Cavigli, E., Danti, G., Galluzzo, M., Granata, V., Pradella, S., Trinci, M., Miele, V., "Role of Chest Imaging in Viral Lung Diseases," *International Journal of Environmental Research and Public Health*, vol. 18, no. 12, p. 6434, 2021. Available: https://doi.org/10.3390/ijerph18126434.

[6] Abdulrahman, A., "PulmoNet: A Novel Deep Learning-Based Pulmonary Diseases Detection Model," *Kaggle*, 2024. Available: https://www.kaggle.com/datasets/abdulahiabdulrahman/pulmonet-dataset.

[7] Rahman, T., "A Multiclass Deep Learning Algorithm for Healthy Lung, Covid-19, and Pneumonia Disease Detection from Chest X-Ray Images," *Kaggle*, 2024. Available: https://www.kaggle.com/datasets/tawsifurrahman/COVID19-radiography-database.

[8] Mooney, P. T., "Contribution to Pulmonary Diseases Diagnostic from X-Ray Images Using Innovative Deep Learning Models," *Kaggle*, 2024. Available: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database.

[9] Blanche, L., "Deep Convolution Neural Network for Respiratory Diseases Detection Using Radiology Images," *Kaggle*, 2023. Available: https://www.kaggle.com/datasets/luisblanche/covidct?select=CT$_COVID$.

[10] Kermany, D., "An Efficient Deep-Learning Model to Diagnose Lung Diseases Using X-Ray Images," *Kaggle*, 2023. Available: https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia.

[11] Abd El-Monem, M., "DarkCVNet: Optimized Pneumonia and COVID-19 Detection Using CXR Images," *Kaggle*, 2022. Available: https://www.kaggle.com/datasets/plameneduardo/sarscov2-ctscan-dataset.

[12] Rahman, T., "Pulmonary Diseases Decision Support System Using Deep Learning Approach," *Kaggle*, 2022. Available: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database.

[13] Elbaz, M., "Diagnosing and Differentiating Viral Pneumonia and COVID-19 Using X-Ray Images," *Kaggle*, 2022. Available: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database.

[14] SIRM Radiology Society, "Analysis of COVID-19 and Pneumonia Detection in Chest X-Ray Images Using Deep Learning," *Italian Society of Medical and Interventional Radiology (SIRM)*, 2021.

[15] Razi, K., "Comparative Analysis of Deep Learning Models for Detection of COVID-19 from Chest X-Ray Images," *Kaggle*, 2020. Available: https://www.kaggle.com/datasets/pranavraikokte/covid19-image-dataset.