Journal of Computer Science

Original Research Paper

A Robust Ensemble Convolutional Neural Networks for Diagnosing Chest Diseases

[1]Mansour Ali Alhlalat, [1]Ahmad Abdel-Aziz Sharieh and [2]Mohammed Belal Al-Zoubi

[1]Department of Computer Science, University of Jordan, Amman, Jordan
[2]Department of Computer Information Systems, University of Jordan, Amman, Jordan

Article history
Received: 26-04-2023
Revised: 27-06-2023
Accepted: 30-08-2023

Corresponding Author:
Mansour Ali Alhlalat
Department of Computer Science, University of Jordan, Amman, Jordan
Email: mansourajarmah@gmail.com

Abstract: Radiologists employ X-ray images to differentiate various chest diseases. Given the intricate and meticulous nature of this diagnostic procedure, the assistance of automated models becomes imperative in detecting and diagnosing diseases from X-ray images. This research paper proposed a novel approach called Ensemble Convolutional Neural Network for Diagnosing Chest Diseases (ECDCNet), aimed at accurately and efficiently diagnosing fifteen different chest diseases through the analysis of X-ray images of the lungs. The ECDCNet model comprised a stack of five CNNs: ResNet152V2, DenseNet121, Inceptionv3, Vogg19, and Wavelet transform-CNN with various architectures and hyper-parameters to enhance the overall prediction performance. The proposed model applied the image segmentation for the lung's region using the U-Net model to localize and focus on the relevant space and facilitate the identification of specific radiological signs such as nodules, opacities, cavities, and consolidation. Furthermore, the study exploited three ensemble CNN strategies: Average voting, majority voting, and a proposed CNN-ensemble strategy called the Weighted Performance Metrics Ensemble Strategy (WPME) to set the weights of the prediction stage. The proposed WPME strategy used four evaluation measures for assessing the importance of each base CNN in the ensemble model, including precision, recall, F1-score, and accuracy, to enhance the prediction of the ensemble model. The proposed ECDCNet model achieved an accuracy of 95.3, 95.8 and 96.1% in the average voting, the majority voting, and the WPME strategy on a collected dataset of 110804 images for fifteen chest diseases. Further, it achieved an accuracy of 97.9, 98.2 and 98.9% in the average voting, the majority voting, and the WPME strategy on another public dataset of 13150 images for three chest diseases.

Keywords: Chest Diseases, Convolutional Neural Network, Deep Learning, Ensemble Learning, Image Processing

Introduction

There are many threats to lung health. For example, the pandemic of Coronavirus (COVID-19) has recently spread terribly and has caused medical and economic problems all over the world. It has become a stumbling block in the continuity of life on this planet (Kaye *et al*., 2021). The respiratory system is highly vulnerable to diseases, with the lung being the largest internal organ in the body and constantly exposed to the external environment (Miller *et al*., 2017).

Respiratory diseases impose a substantial economic burden on healthcare systems due to their associated costs. A notable example is the treatment of asthma, which incurs significant expenses. According to recent research by Lee *et al*. (2021), the annual cost of asthma treatment in the United States of America alone amounts to approximately 18 billion dollars.

X-ray imaging is a valuable resource in radiology, aiding in differentiating various chest diseases (Ait Nasser and Akhloufi, 2023). However, the diagnostic process can be complex and challenging for radiologists. As a result, there is a growing need for assistance in detecting and diagnosing diseases from X-ray images, including conditions such as pneumonia, atelectasis, cardiac hypertrophy, effusions, nodules, and COVID-19 (Erdaw and Tachbele, 2021). Consequently, there is a significant demand for fast,

effective, and accurate models that diagnose chest diseases. Over the past two decades, deep neural networks made remarkable strides in object detection. In particular, object detection in deep neural networks consists of two main types: The two-stage approach and the one-stage approach. The two-stage demonstrated superior accuracy, while the one-stage showed greater efficiency (Zhang *et al*., 2018). In the detection of chest diseases in X-ray images, it becomes crucial to achieve both high efficiency and accuracy. Therefore, there is a need to enhance the accuracy of the one-stage or improve the efficiency of the two-stage when utilizing deep neural networks for such detection processes in X-ray images.

This research proposed a novel contribution to diagnosing chest diseases through an improved ensemble model called Ensemble Convolutional Neural Networks for Diagnosing Chest diseases (ECDCNet), including cardiomegaly, emphysema, edema, hernia, pneumothorax, effusion, mass, fibrosis, atelectasis, consolidation, pleural thickening, nodule, pneumonia, infiltration, and Covid-19. The ECDCNet model consists of five CNNs: ResNet152V2, DenseNet121, Inceptionv3, Vogg19, and Wavelet transform-CNN trained using two different approaches: Image-based training and segmented-based training. Further, the study presented a new conventional ensemble learning strategy called the Weighted Performance Metrics Ensemble strategy (WPME) to improve prediction accuracy. The WPME incorporates precision score, recall score, F1-score, and accuracy as performance metrics to optimize the predictive capability of the ECDCNet model.

The structure of this research paper is as follows: The first section provided an overview of chest diseases, object detection using Convolutional Neural Networks (CNNs), ensemble methods in deep learning, and relevant prior studies in diagnosing chest diseases. In the subsequent section, the components of the ECDCNet model were described, along with details about the datasets used and the evaluation metrics employed. The third section focused on the analysis of obtained results. Subsequently, the fourth section discussed the obtained results and summarized the findings. The fifth section concluded the study and proposed potential avenues for future research. Finally, the paper included acknowledgments, funding information, author contributions, and an ethics statement.

Background and Related Work

This section explored prior research on object detection of chest diseases using CNNs. It provided an overview of the critical components utilized in CNNs, explored ensemble strategies in deep learning, and discussed the application of wavelet transformations in this context.

Chest Diseases

A chest disease is a problem in the chest that blocks it from properly working. Chest diseases pose a significant threat to people's human health, often leading to severe consequences and loss of life (Cohen *et al*., 2022). Ongoing medical advancements and research endeavors aim to deepen our understanding of these diseases and develop effective treatment strategies. Early detection, timely interventions, and appropriate management are critical in improving outcomes for affected persons with chest diseases. Radiologists commonly rely on X-ray imaging to identify and differentiate various chest diseases by detecting anomalies within the images. Anomalies typically manifest as texture abnormalities, focal abnormalities, or abnormal shapes. Texture abnormalities are identified through widespread changes in appearance and structure, while focal abnormalities indicate localized changes in density. Also, the deviation in the outline of the anatomy refers to abnormal shapes. (van Ginneken *et al*., 2009). Notable chest diseases include atelectasis, cardiomegaly, consolidation, infiltration, nodule, mass, pulmonary fibrosis, pneumothorax, pleural thickening, pleural effusion, and Covid-19 (Wang *et al*., 2017; Shih *et al*., 2019; Lakhani *et al*., 2023). Figure 1 common types of chest diseases observed in X-ray images.

X-ray images emerged as valuable resources for radiology, providing significant support in diagnosing respiratory diseases, such as aortic enlargement, atelectasis, cardiomegaly, calcification, and others (Ait Nasser and Akhloufi, 2023). However, the diagnostic process is intricate and requires the expertise of radiologists who meticulously analyze these images. This study proposed an ensemble of CNNs that helps to solve these challenges.

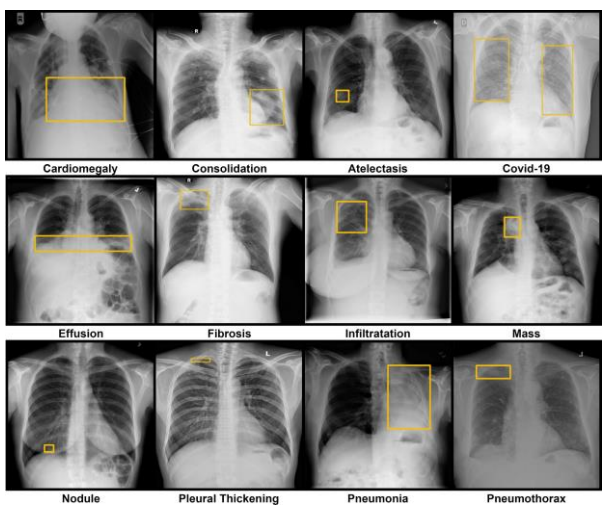

Fig. 1: Common chest diseases observed in the X-ray images 19 (Wang *et al*., 2017; Shih *et al*., 2019; Lakhani *et al*., 2023)

Convolutional Neural Networks

A Convolutional Neural Network (CNN) is a machine learning algorithm utilized for object detection in input images to classify or predict these images. By detecting image features, the CNN can recognize the image or object (Sapijaszko and Mikhael, 2018). The CNN architecture consists of three layers: Convolutional layers, pooling layers, and fully-connected layers.

The convolutional layer can extract and identify various features, corners, objects, and edges from the input images by applying the convolution operation between the input image and a kernel of specific dimensions (K × K). The pooling layer reduces computational costs by minimizing the features extracted from the convolutional layers. The fully connected layer can predict the most appropriate class for the input image.

Further, a dropout layer and activation function are incorporated into the design of the CNN. The dropout layer enhances the model's performance by selectively deactivating the neurons in the training process to mitigate the risk of overfitting. The activation function plays a crucial role in determining whether a neuron should be activated or not activated during the prediction phase (Yamashita *et al*., 2018).

He *et al*. (2016) introduced a framework known as ResNet to address the degradation in deep neural networks. The ResNet framework utilizes the concept of residual mapping and introduces a specialized building block called the residual block. By incorporating skip connections, the ResNet model allowed the flow of information from earlier layers to subsequent levels. This architecture was composed of multiple stacked residual blocks. Different variations of ResNet, such as those with 18, 34, 50, 101, or 152 layers, were widely adopted in various applications.

Huang *et al*. (2017) proposed the Dense convolutional Network (DenseNet), a CNN designed to address the vanishing gradient problem and improve accuracy. DenseNet utilizes dense connections between layers, where each layer receives inputs from all preceding layers. This approach enhanced computational and memory efficiency. The DenseNet architecture consists of multiple dense blocks containing several convolutional layers. The primary components of DenseNet include composition, bottleneck, transition, and classification layers. The composition layers consist of a BatchNorm layer, a ReLU layer, and a 3×3 convolutional layer. The bottleneck layers include two batch norm layers, a ReLU layer, a 1×1 convolutional layer, a ReLU layer, and a 3×3 convolutional layer. Transition layers involve a 1×1 convolutional layer followed by a 2×2 pooling layer with a stride of 2. The classification layer incorporates a 7×7 global average pooling, a fully connected layer, and a softmax function.

Szegedy *et al*. (2015) proposed a CNN (InceptionV1) to enhance the utilization of network computing resources. The InceptionV1 model consists of a series of inception modules, each consisting of convolution layers with kernel sizes of 1×1, 3×3, or 5×5. These modules also include max-pooling layers and concatenation layers. In total, the InceptionV1 model is composed of 22 layers.

Szegedy *et al*. (2016) modified the InceptionV1 architectures by proposing the InceptionV2 and InceptionV3 models to enhance the computational power by using many techniques such as factorized convolutions, dimension reduction, parallelized computations, and regularization. The InceptionV2 was designed in three versions of inception modules based on three different principles: Smaller convolutions factorization, asymmetric convolutions factorization, and the inclusion of auxiliary classifiers. In inceptionv3, the authors integrated the upgrades in InceptionV2 and used the batch norm, the label smoothing, the factorization for the 7×7 convolution, and the RMS prop optimizer.

Simonyan and Zisserman (2014) designed a very deep convolutional network (VGG16) that achieves 92.7% accuracy in the ImageNet dataset. This network consists of thirteen convolutional layers, five max-pooling layers, and three dense layers. VGG16 exchanges the big kernel size with many successive 3×3 kernel sizes. The authors proposed another version of this network called VGG19 which is the same as VGG16 with supports 19 convolutional layers.

Ronneberger *et al*. (2015) proposed the U-Net for semantic segmentation. The U-Net consists of contraction blocks and expansive blocks. The contraction block consists of two 3×3 convolutions followed by a ReLU and max pooling. Also, the expansive block consists of a transpose convolution of size 3×3 with stride 2, a concatenation with the correspondingly max pooling outputs from the contracting block, and two 3×3 convolutions, each followed by a ReLU. The output from the last expansive block is applied to a 1×1 convolution and sigmoid function.

Ensemble Strategies in Deep Learning

In deep learning, ensemble learning merges multiple CNN models to achieve a consensus in predictions. Ensemble learning is utilized in various contexts, including improving classification performance, estimating confidence levels, obtaining optimal models, and training with limited or excessive data (Mohammed and Kora, 2023).

The common strategies employed in ensemble learning include bagging, boosting, stacking, a mixture of experts, majority voting, probability averaging, and weighted probability averaging (Ganaie *et al*., 2022).

The bagging strategy divides the dataset into random subsamples, known as bootstrap sampling. This strategy

selects models with low bias and high variance. The boosting strategy entails feeding the first model with the entire dataset while subsequent models receive misclassified data from the preceding models. This iterative process aims to reduce bias in predictions (Ganaie *et al*., 2022).

The stacking strategy operated similarly to the bagging, but the stacking divides the datasets into random subsamples. However, the outputs from all models were fed into a meta-classifier to generate the final prediction. In contrast, the mixture of experts' strategies involved putting the entire dataset to all models but assigning weights to each model through a trainable model known as the "gating network" (Ganaie *et al*., 2022). The majority voting strategy tallied the predictions for each class and selected the one has the highest prediction count. On the other hand, the probability averaging strategy computed probability scores for each model and averaged them across all models. Weighted probability averaging functioned similarly but incorporated a weighted average of the models based on their significance (Kundu *et al*., 2021).

This study proposed a novel weighted probability averaging strategy that utilized four evaluation measures: Precision score, recall score, F1 score, and accuracy. These measures were employed based on the approach deployed in the study by Kundu *et al*. (2021) but with a novel strategy aimed at enhancing the final prediction of the ensemble model.

Wavelet Transform

This section clarified three fundamental concepts:

- The first concept is the Spatial Domain (SD), which refers to the arrangement of pixels in an image and the representation of the image's content (Al Jumah, 2013). The SD domain allows control over individual pixel values and their neighbors. The SD can perform many image processing tasks, such as noise removal, edge detection, and image restoration
- The second concept is the Frequency Domain (FD), FD in image processing refers to the rate of change in the pixels of an image. In the FD, the concern is with which rates the pixel values are changing in the spatial domain. So, applying filters to modify or enhance an image in the frequency domain is faster than in the spatial domain (Al Jumah, 2013)
- The third concept entails Wavelet Transformation (WT), a mathematical technique that centers on both time and frequency domains surrounding a specific point. It facilitates the decomposition of an image into high-frequency components (approximations) and low-frequency components (details) (Parida and Bhoi, 2017)

The WT achieves superior frequency resolution for details and high temporal resolution for approximations. Additionally, the best approach to handling the essential regions, which hold the significant data source in an image, is to identify those abrupt changes in contrast by combining time and frequency domains through wavelet techniques to solve the challenges posed by analyzing and manipulating data in these regions by leveraging the capabilities of wavelet transformations to capture both temporal and frequency information in the data. By applying wavelet techniques, researchers can extract meaningful features and representations from the data, enabling more accurate analysis and manipulation of these challenging regions. Parida and Bhoi (2017). Figure 2 shows the first-level decomposition of the WT in the X-ray image. The image passes through two filters: A high-pass and a low-pass, to decompose into high-frequency and low-frequency components that achieve four sub-images (low-low, low-high, high-low, and high-high) at each level, showcasing the comprehensive nature of the wavelet transformation process.

Related Work

This section introduced an overview of research studies in the CNNs for chest disease diagnosis. The first part of this section highlighted the research studies that explored the use of CNNs as standalone models in chest disease diagnosis. Building upon the advancements in CNN-based approaches, the second part of this section discussed research studies that utilized ensemble CNNs in chest disease diagnosis. Moreover, the third part of this section introduced research studies that utilized lung segmentation techniques in chest disease diagnosis. Examining these studies identified the strengths and limitations of existing approaches to address the challenges associated with chest disease diagnosis.

CNNs in Chest Disease Diagnosing

CNNs emerged as powerful machine learning models for image-related tasks, garnering significant attention in chest disease diagnosis (Taye, 2023; Sarvamangala and Kulkarni, 2022).

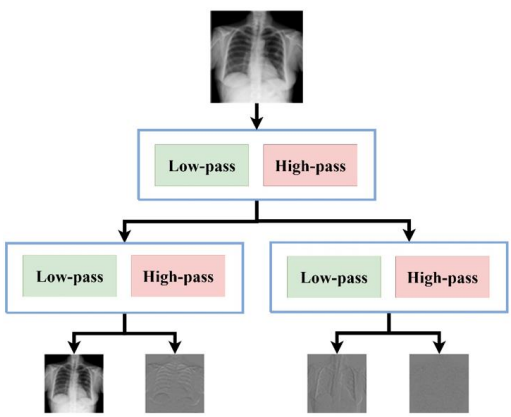

Fig. 2: Wavelet transform of the X-ray image

Table 1: Some studies of lung disease diagnosing using CNNs

Study	Model	No. of classes	Accuracy
Singh *et al*. (2023)	CNN: Quaternion residual attention network	2	94.53
Mann *et al*. (2023)	DenseNet121	14	90.00
	ResNet50	14	91.00
Malik *et al*. (2023)	Vgg-19	6	95.61
	ResNet-50	6	96.15
	Inception v3	6	95.16
Sharma *et al*. (2022a)	VGG19	3	91.67
	ResNet50	3	90.33
	MobileNetV2	3	76.00
	SqueezeNet	3	97.33
Pramanik *et al*. (2022)	DenseNet121	3	96.96
	ResNet101	3	88.01
	VGG-19	3	96.03
	InceptionV3	3	95.10
Prakash *et al*. (2023)	ResNet101V2	4	95.00
	DenseNet121	4	95.00
Visuña *et al*. (2022)	VGG19	4	96.00
	InceptionV3	4	92.00
Dey *et al*. (2021)	VGG19	2	97.94
Almezhghwi *et al*. (2021)	AlexNet	13	94.00
	VGG16	13	95.00

Their remarkable ability to enhance generalization performance compared to other machine learning models propelled CNNs to the forefront of disease diagnosis research. In particular, several studies utilized CNNs to improve the performance of chest disease diagnosis in X-ray images, taking advantage of their capability to extract intricate features from medical images. Table 1 highlights the contributions and achievements of these studies. The demonstrated achievements of these studies establish a solid foundation for further advancements in chest disease diagnosis using CNNs.

Ensemble CNNs in Chest Disease Diagnosing

Ensemble CNNs garnered significant attention in disease diagnosing due to their ability to improve generalization performance compared to individual CNNs and other machine learning approaches. Several studies explored ensemble CNNs, which combine multiple CNN models to enhance performance and improve generalization.

One notable study by Kotei and Thirunavukarasu (2022) proposed a stacked ensemble algorithm that classified the X-ray images as tuberculosis-positive or negative. The proposed model achieved an accuracy of 98.89%, a sensitivity of 98.38%, and a specificity of 98.70%. The study used different CNN models. The proposed model achieved 98.58, 93.10, and 96.50% accuracy, IoU, and dice coefficient, respectively. These results demonstrated that ensemble learning with segmented CXR images outperformed unsegmented CXR images.

In a similar El Lel *et al*. (2023) proposed eleven ensemble CNNs (etc., random forest classifier, logistic regression, k-neighbors classifier, and majority voting) consisting of six CNNs and a classifier layer to classify the input image into COVID-19 or normal. Through rigorous evaluation, the logistic regression ensemble model emerged as the most effective in terms of COVID-19 detection accuracy, achieving an impressive accuracy rate of 96.29%.

Prakash *et al*. (2023) proposed a novel stacked ensemble learning approach for pediatric pneumonia classification. The proposed method achieved remarkable results with an accuracy of 98.3%, precision of 99.29%, recall of 98.36%, F1-score of 98.83%, and an impressive AUC score of 98.24%. These performance metrics surpass those achieved by other classifiers used in the study.

Kalaivani and Seetharaman (2022) developed a three-stage ensemble-boosted CNN model for diagnosing chest diseases: COVID-19, TB, and BP. In the initial step, the study used the ResUNet model to improve the performance of processing the CXR datasets. Then, it employed an independent CNN to extract features from the images in the training dataset. The proposed model achieved an accuracy of 99.35% and a recall of 98%.

Agrawal *et al*. (2022) proposed an ensemble of deep CNNs models termed RGDSoft that showed a good result on different datasets. The RGDSoft used a pre-trained classifier: DenseNet-121, GoogLeNet, and ResNet-18 to detect pneumonia and COVID-19 infections in the X-ray images. The RGDSoft model achieved 98.1, 98.8, and 98.8% for accuracy, precision, and recall on the Kermany dataset. Also, the model achieved 94, 96, and 96% accuracy, precision, and recall on the COVID-19 dataset.

Visuña *et al*. (2022) proposed two distinct CNN-ensemble Voting and Stacking models. The proposed

models were designed and evaluated for chest disease diagnosis. These ensembles combined predictions from six CNN models: VGG16, VGG19, DenseNet121, ResNet101V2, InceptionV3, and CheXnet. The proposed models achieved impressive accuracy rates of 99% for the voting ensemble and 98% for the stacking ensemble. The proposed models were validated using a diverse dataset: Viral pneumonia, COVID-19, and Tuberculosis categories.

These aligned findings by previous research studies demonstrated that ensemble CNN models exhibit superior performance and generalization capabilities compared to individual CNN models when dealing with various chest diseases. The ability of ensemble CNNs to leverage the diversity of multiple models enhances their robustness and overall accuracy in disease diagnosis. The robustness and reliability demonstrated by the ensemble model make it suitable for deployment in assisting radiologists and physicians in diagnosing pediatric pneumonia, thereby enhancing the accuracy and efficiency of disease detection.

However, this study observed three main challenges in previous research studies (Bhatt and Shah, 2023; El Lel *et al*., 2023; Mabrouk *et al*., 2022; Visuña *et al*., 2022; Win *et al*., 2021; Deb and Jha, 2020). The first challenge is the availability of diverse and sufficient data is crucial for the success of ensemble CNNs. The second challenge is that ensemble CNN training requires high computation resources. The last challenge lies in consensus among multiple models to attain a unified decision or prediction.

Despite these challenges, this study handled and developed a robust and accurate ensemble CNN model that can diagnose chest diseases in X-ray Images.

Lung Segmentation Using U-Net

Lung segmentation is a crucial component that helps to identify lung regions (Teixeira *et al*., 2021). This segmentation technique enables machine learning models to focus specifically on the areas of interest related to lung abnormalities, leading to increased efficiency during both the training and prediction stages. Deep learning, particularly CNNs such as the U-Net model, emerged as the preferred approach for image segmentation. Several research studies employed lung segmentation techniques to augment the performance of CNNs in lung disease diagnosis (Arvind *et al*., 2023; Chang *et al*., 2022; Kumarasinghe *et al*., 2022; Cao and Zhao, 2021; Lee *et al*., 2022; Kim *et al*., 2022; Schultheiss *et al*., 2021; Bosdelekidis and Ioakeimidis, 2020). These studies unequivocally demonstrated the pivotal role of lung segmentation in improving the accuracy and automation of chest disease diagnosis. Significantly, the U-Net architecture emerged as an impressive choice in this context (Sogancioglu *et al*., 2020; Kumarasinghe *et al*., 2022; Sharma *et al*., 2022b; Gadgil *et al*., 2021). The U-Net architecture offers several advantages, including its efficient inference time, facilitating rapid analysis of lung images (Kumarasinghe *et al*., 2022; Gadgil *et al*., 2021). This characteristic renders it highly suitable for applications requiring real-time or near-real-time results, such as lung disease diagnosis. Moreover, the U-Net architecture's efficacy in accurately delineating lung boundaries established it as a favored option among researchers in the field (Zhao *et al*., 2022; Gadgil *et al*., 2021; Jaszcz *et al*., 2022).

In recent studies, Bassi and Attux (2022) proposed segmentation and classification methods using deep neural networks to identify COVID-19 and pneumonia diseases in X-ray images. Specifically, these studies employed the U-Net architecture for lung segmentation and the DenseNet201 model for disease classification. The authors achieved notable performance improvements. The DenseNet201 model achieved an impressive Area Under the Curve (AUC) value of 0.917 when combined with lung segmentation, compared to 0.906 without segmentation.

In their study, Nillmani *et al*. (2022) presented a comprehensive investigation into the detection of COVID-19 in X-ray images using CNNs and lung segmentation techniques. The researchers developed sixteen distinct classification models, combining the power of CNN architectures with lung segmentation. The Authors employed two models: UNet and UNet+, for the lung segmentation task. The UNet model emerged as the top-performing segmentation model, achieving an impressive accuracy of 96.35%. Moreover, the UNet with Xception model achieved an accuracy of 97.45%.

Sogancioglu *et al.* (2020) introduced a novel model for cardiomegaly detection on X-ray images. This model utilized two deep-learning approaches, anatomical segmentation that deployed the U-net model and image-level classification. The segmentation-based model demonstrated exceptional performance, achieving an impressive Area Under the Curve (AUC) value of 0.977. Similarly, the image-level classification model achieved an AUC of 0.941. Further, the Authors demonstrated that the segmentation-based model outperformed the image-level classification model, requiring significantly fewer annotated chest X-rays. The segmentation-based model showcased a remarkable advantage by achieving substantially better performance with only $1/100^{th}$ of the annotated data, underscoring its potential for cost-effective and resource-efficient diagnosis of cardiomegaly.

Materials and Methods

This study introduced the ECDCNet model, a novel approach for chest disease diagnosis in X-ray images. The proposed ECDCNet model utilized six distinct CNNs to enhance classification performance. These base CNNs include ResNet105V2, DenseNet121, VGG19, InceptionV3, Wavelet-CNN and UNet. The first five CNNs were employed for chest disease diagnosing. These

CNNs predicted separately to get confidence prediction results by obtaining varied features of every image from each model to enhance the classification performance. Also, the U-Net CNN was used for lung segmentation in the input X-ray image. This segmentation technique was utilized to extract lung-related regions of interest, enabling more accurate and precise disease prediction.

The ECDCNet model has presented a comprehensive framework for chest disease diagnosing through robust six-based CNNs and lung segmentation with U-Net. The ECDCNet contributed to improved classification performance and provided valuable insights for accurate diagnosis and treatment planning in the field of chest diseases.

This section introduced the study motivation, the architecture of the proposed ECDCNet model, the proposed strategy for the ensemble model, the training and prediction process for the proposed model, the utilized data, and the applied data preprocessing techniques.

Motivation

In lung disease diagnosis, many studies predominantly focused on binary classification, with limited exploration of multiclass classification. Also, while some studies achieved good performance in binary classification (Singh *et al*., 2023; Dey *et al*., 2021), the multiclass classification often resulted in lower performance (Mann *et al*., 2023; Almezhghwi *et al*., 2021). This study addressed this gap by proposing a model capable of classifying sixteen different classes: To diagnose sixteen different classes: Cardiomegaly, emphysema, edema, hernia, pneumothorax, effusion, mass, fibrosis, atelectasis, consolidation, pleural thickening, nodule, pneumonia, infiltration, Covid-19 and no finding, thereby expanding diagnostic capabilities and efficiency.

Image-based CNNs encounter difficulties detecting subtle details or distinct objects of interest in an image (Bassi and Attux, 2022; Sharma *et al*., 2022b; Sogancioglu *et al*., 2020). However, segmentation techniques offered a solution by enabling machine learning models to concentrate on the regions associated with lung abnormalities. This study deployed a segmentation technique utilizing the robust U-Net model to enhance the efficiency during the training and prediction stages.

Ensemble models gained significant attention in disease diagnosis due to their ability to enhance generalization performance compared to individual CNNs and other machine-learning approaches (El Lel *et al*., 2023; Prakash *et al*., 2023; Mabrouk *et al*., 2022). However, ensemble models face unique challenges in this field (Bhatt and Shah, 2023; El Lel *et al*., 2023; Mabrouk *et al*., 2022; Visuña *et al*., 2022; Win *et al*., 2021; Deb and Jha, 2020). The study overcame these challenges by introducing a novel procedure consisting of the following key points:

- A crucial challenge in ensemble models is achieving consensus among multiple models to reach a unified decision or prediction. The study overcomes these by proposing WPME strategy leverages a combination of four evaluation measures: Precision, recall, F1-score, and accuracy. By incorporating multiple evaluation measures, the WPME enhanced the prediction capabilities of the ensemble model, resulting in more reliable and accurate results
- The availability of diverse and sufficient data is essential for the success of ensemble CNNs. This proposed study handled this challenge by collecting data from five different resources to create a comprehensive dataset with a wide range of cases. Also, the study employed data augmentation techniques to enhance the variability of the available data, which plays an essential role in increasing the robustness and generalization capabilities of the model
- The performance of ensemble CNNs is inherently dependent on the performance of the base CNNs. In overcoming this challenge, the study effectively utilized a combination of robust CNNs that showed their capability in lung diseases diagnosing: ResNet152V2, DenseNet121, Inceptionv3, Vogg19, and Wavelet transform-CNN integrating them to achieve consensus and ultimately reach a unified decision or prediction
- The training of the ensemble CNNs required high computation resources. In overcoming this challenge, the studies implemented the early stopping criteria using Receiver Operating Characteristic (ROC) analysis. This approach allowed us to monitor the model's performance during training and stop the training process when further iterations did not significantly improve the model's performance

This study contributed to lung disease diagnosis with a comprehensive and accurate ensemble model that h utilized several concepts: Segmentation techniques, ensemble models, and the WPME strategy to enhance the accuracy and reliability of lung disease classification, ultimately benefiting patients and medical professionals in their diagnostic endeavors.

Datasets Description

The U-Net was trained and evaluated using the Chest X-ray dataset for lung segmentation dataset (Danilov *et al*., 2022), which consists of 6,810 chest images with their binary masks for the lungs region. This

data was collected from three datasets which are Shenzhen, Darwin, and Montgomery with 566, 6106, and 139 images, respectively. The main dataset was divided into three data subsets with 64, 16, and 20% for training, validating, and testing, respectively.

To train and evaluate the ECDCNet model, a dataset was created by collecting images from three datasets termed ECD1, which are Chest X-ray14 Dataset (CXD) (Wang *et al*., 2017; Shih *et al*., 2019; Lakhani *et al*., 2023), COVID-19 Radiography Database (CRD) (Rahman *et al*., 2021; Chowdhury *et al*., 2020) and Dataset of RSNA Pneumonia Detection (RPDD) (Shih *et al*., 2019). (CXD) (CRD) (RPDD).

The dataset of RSNA pneumonia Detection (RBD) consists of 30,000 chest X-ray images, classifying the sample as to whether it has pneumonia or not. Also, chest X-ray 14 includes 112,120 frontal-view X-ray images of 30,805 patients with the fourteen diseases, it contains images for most chest diseases. Further, the COVID-19 radiography database includes 3616 COVID-19 positive cases along with 10,192 normal, 6012 lung opacity, and 1345 viral pneumonia images database. The Collected Dataset (CD) consists of 110804 images for fifteen diseases. Also, the ECDCNet model was trained and evaluated using A second collected dataset from two public datasets which are chest X-ray images (pneumonia) from Kaggle (Kermany *et al*., 2018), SIIM-FISABIO-RSNA COVID-19 4295 images for COVID-19 (Lakhani *et al*., 2023). Consiste of Covid-19 (4275), pneumonia (3875), normal (5000) images. This dataset is termed ECD2.

Base CNNs

This study employed six base CNNs: VGG19, InceptionV3, ResNet152V2, DenseNet121, and wavelet-CNN. Each CNN operated independently, focusing on extracting distinct features from the input images. After the base CNNs generated his prediction, the ensemble models utilized the combined predictions from the base models to generate the final diagnostic class.

VGG19 consisted of 19 layers: Two 3×3 convolution with stride 1 and 64 output channels, MaxPool layer of 2×2 pixel windows with stride 2, two 3×3 convolution layers with stride 1 and 128 output channels, MaxPool layer of 2×2 pixel windows with stride 2, four 3×3 convolution layers with stride 1 and 256 output channels, MaxPool layer of 2×2 pixel windows with stride 2, four 3×3 convolution layers with stride 1 and 512 output channels, MaxPool layer of 2×2 pixel windows with stride 2, four 3×3 convolution layers of size with stride 1 and 512 output channels, MaxPool layer of 2×2 pixel windows with stride 2, two fully connected layers of size 4096 and other fully connected layers with 1000 channels.

InceptionV3 consisted of 42 layers: 3×3 Convolution layer with stride 2 and 32 output channels, 3×3 Convolution layer with stride 1 and 32 output channels, 3×3 Padded Convolution with stride 1 and 64 output channels, MaxPool layer of 2×2 pixel windows with stride 2, 3×3 Convolution layer of size 3×3 with stride 1 and 80 output channels, 3×3 Convolution layer with stride 2 and 192 output channels, 3×3 Convolution layer of size 3×3 with stride 1 and 288 output channels, three Inception Module1, five Inception Module 2, two Inception Module 3, MaxPool layer of 2×2 pixel windows with stride 2. Linear and softmax classifier.

DenseNet121 consisted of 120 layers: A 7×7 convolution layer with stride 2 and 112 output channels, a MaxPool layer of 3×3-pixel windows with stride 2, Dense Block 1 with 56 output channels, transition layer 1: (A 1×1 convolution layer with stride 1 and 56 output channels. Followed by an average pool layer of size 2×2 with stride 2 and 28 output channels), Dense Block 2 with 28 output channels, transition layer 2: (A 1×1 convolution layer with stride 1 and 28 output channels. Followed by an average pool layer of size 2×2 with stride 2 and 14 output channels), Dense Block 3, transition layer 3: (A 1×1 convolution layer with stride 1 and 14 output channels. followed by an average pool layer of size 2×2 with stride 2 and 7 output channels), Dense Block (4) with seven output channels and classification layer: (global average pool Layer with size 7×7 and 1000D fully connected layer-softmax).

ResNet152V2 consisted of 152 layers: Conv2 with 56 output channels: (MaxPool layer of 3×3 pixel windows with stride 2, followed by three (1×1 convolution layer with stride 1, 3×3 convolution layer with stride 1 and 1×1 convolution layer of size 1×1 with stride 1)), Conv3 with 28 output channels: (Eight (1×1 convolution layer with stride 1, 3×3 convolution layer 3×3 with stride 1 and 1×1 convolution layer of size 1×1 with stride 1)), Conv4 with 14 output channels: (Thirty-six (1×1 convolution layer with stride 1, 3×3 convolution layer with stride 1 and 1×1 convolution layer of size 1×1 with stride 1)), Conv5 with seven output channels: (A three (1×1 convolution layer with stride 1, 3×3 convolution layer of size with stride 1 and 1×1 convolution layer of size 1×1 with stride 1)) and 7×7 average pool layer and 1000 D fully connected layer.

Wavelet-CNN consisted of two branches: The first branch is: Three 3×3 convolutional layers with stride 1, a global average pooling layer, a dropout layer, and a flattened layer. Further, the second branch receives the output images from the wavelet decompositions: The first level decomposed image is directed to two 3×3 convolution layers with stride 1. The output result from the previous step is concatenated to the second-level decomposed image and fed to two 3×3 convolution layers of size 3×3 with stride 1. The result is concatenated to the

third-level decomposed image. Also, it is assigned to a global average pooling and flattening layers. Finally, the outputs from the first and second branches are concatenated and fed to three dense layers.

The U-Net consisted of five contraction blocks and four expansive blocks. Each contraction block consists of a sequence of operations, including a 3×3 convolution layer, a ReLU activation function, dropout regularization, another 3×3 convolution layer, ReLU activation, and MaxPooling. Conversely, each expansive block performs a series of operations t, including a 3×3 transpose convolution and stride 2, followed by a 3×3 convolution layer, ReLU activation, dropout, another 3×3 convolution layer, and ReLU activation. Figure 3 shows the U-Net model.

Ensemble Models

Ensemble learning is a method that combines the predictions of multiple models to obtain a consensus. However, many ensemble learning strategies determine the weights of the classifiers based on their accuracy or through empirical methods. These strategies may not yield optimal predictions in all situations, such as when dealing with imbalanced datasets. This study employed three strategies: Majority voting, weighted majority voting, and a newly proposed strategy called Weighted Performance Metrics Ensemble (WPME). The WPME strategy utilizes four evaluation measures, precision, recall, F1-score, and accuracy, to improve the prediction of the ensemble model.

The majority voting approach selects the class with the highest frequency predicted by the base CNNs in the ECDCNet model. Equation 1 shows the operation of majority voting:

$$Prdection(x) = Mode(pre(x)_1, pre(x)_2, \ldots, pre(x)_N) \quad (1)$$

where, *Prediction(x)* is the final predicted class label of the input image *x*, $pre(x)_N$ is the predicted class label of the image *x* in the N^{th} base CNN and the *Mode* calculates the class label with the highest occurrence among all CNNs.

The study utilized a weighted majority ensemble to improve the model prediction performance. In the training process, the classifier updates the weights of each base CNN based on the prediction accuracy that is achieved, with higher weights assigned to a more accurate base CNN. During prediction, the classifier computes the weighted sum of the base CNNs' predictions for each class label by multiplying each base CNN's prediction by its corresponding weight and summing the results. Finally, the classifier selects the class label with the highest weighted sum as the final prediction.

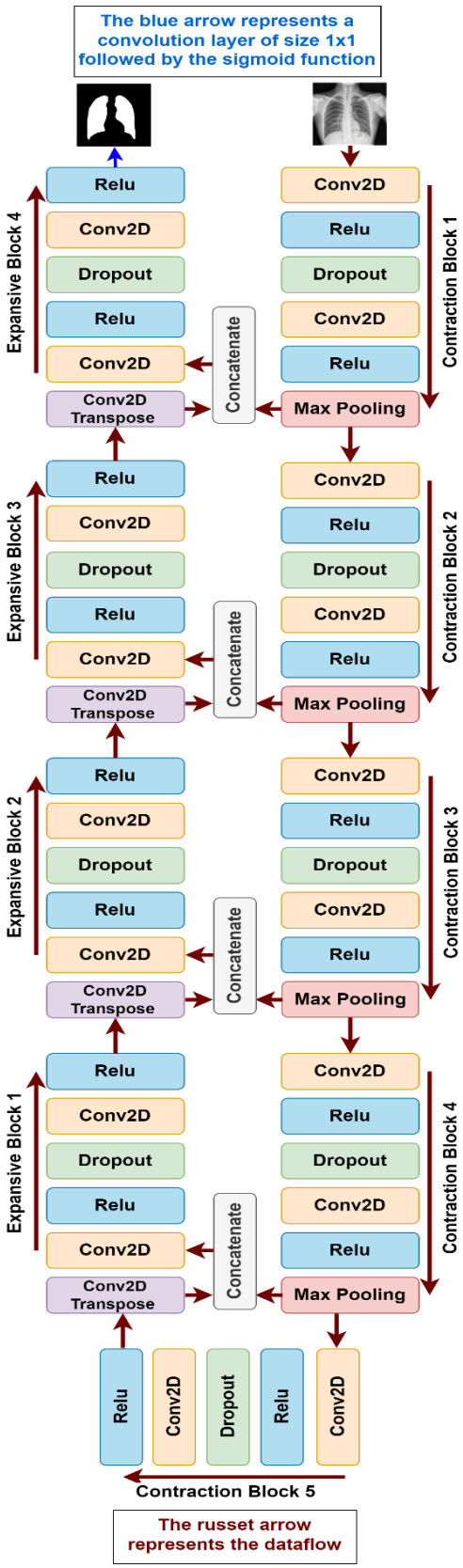

Fig. 3: The U-Net model

Further, the study proposed a WPME ensemble to improve the model prediction performance. In the training process, the classifier updates the weights of each base CNN based on the prediction's four metrics recall score, F1 scores, and accuracy.

During prediction, the classifier follows the following steps:

Step 1: Collect the scores obtained from each base CNN for a specific metric and apply the softmax function to those scores. Repeat this process for all metrics (precision, recall, F1-score, and accuracy). Using Eqs. 2-5:

$$PrS = softmax(PrV) \tag{2}$$

where, $PrV = [Pr_1, Pr_2, ..., Pr_n]$ represents the precision values of all base CNNs and $PrS = [PrS_1, PrS_2, ..., PrS_n]$ represents the normalized precision values:

$$RCS = softmax(RCV) \tag{3}$$

where, $RCV = [R_1, R_2, ..., R_n]$ represents the recall values of all base CNNs and $RS = [RS_1, RS_2, ..., RS_n]$ represents the normalized recall value:

$$F1S = softmax(F1V) \tag{4}$$

where, $F1V = [F1_1, F1_2, ..., F1_n]$ represents the F1-score values of all base CNNs and $F1S = [F1S_1, F1S_2, ..., F1S_n]$ represents the normalized *F1* core values:

$$AccS = softmax(AccV) \tag{5}$$

where, $AccV = [Acc_1, Acc_2, ..., Acc_n]$ represents the accuracy values of all base CNNs and $AccS = [AccS_1, AccS_2, ..., AccS_n]$ represents the normalized accuracy values.

Step 2: For each base CNN, apply the softmax function for its normalized scores from these vectors *PrS*, *RS*, *F1S*, and *AccS* using Eq. 6:

$$M_x = softmax(PrS_x, RCS_x, F1S_x, AccS_x) \tag{6}$$

where, PrS_x, RS_x, $F1S_x$, and $AccS_x$ are the modified normalized vectors precision, recall score, F1-score, and accuracy score of the x^{th} base CNNs, respectively. Where, $M_x = [PrM_x, RM_x, F1M_x, AccM_x]$ represents the scores after the second normalization for the x^{th} base CNN.

Step 3: For each base CNN, calculate its weight by adding its normalized scores using Eq. 7:

$$W_x = C_0 \times PrM_x + C_1 \times RM_x + C_2 \times F1M_x + C_3 \times AccM_x \tag{7}$$

where, W_x is the weight of the x^{th} model, C_0, C_1, C_2, *and* C_3 are constant (a value between 0 and 1) assigned to each metric based on its significance.

Step 4: Computes the weighted sum of the base CNNs' predictions for each class label by multiplying each base CNN's prediction by its corresponding weight and summing the results. Finally, the classifier selects the class label with the highest weighted sum as the final prediction

The Training Process

The training process for ECDCNet involves the following steps:

Step 1: A large dataset of X-ray chest mages, including different diseases, was gathered. A dataset was created by collecting X-ray images from four public datasets

Step 2: Several preprocessing techniques were applied to the collected images to ensure their suitability for the lung disease diagnosis task

- Images were converted from DICOM to PNG for easier handling and compatibility across different datasets
- Noise removal and size standardization techniques were applied to ensure that all input images were in a suitable format and of optimal quality
- The mage segmentation was applied for the lung's region extraction purpose to localize and focus on the relevant space and facilitate the identification of specific radiological signs such as nodules, opacities, cavities, and consolidation
- Image augmentation techniques were employed during each epoch of the training process. Specifically, the input images randomly get transformations: Shifting, zooming, and rotation. This step introduced variations and increased the diversity of training samples. By doing so, the model became more capable of generalizing and making accurate predictions on unseen data

It is important to note that the applied preprocessing approach is conservative, aiming to minimize information loss. The steps involved essential operations while avoiding excessive filtering or aggressive modifications that could potentially remove crucial diagnostic features. This balance between preprocessing and preserving relevant information ensures that the model is robust enough to make accurate and reliable predictions for lung disease diagnosis.

Step 3: Divide the preprocessed datasets into three sets: Training, validation, and testing, following the configuration presented in Table 3. The training set was utilized to update the model's weights, while the validation set played a crucial role in monitoring the model's performance throughout the training process. Finally, the testing set was employed to assess the overall performance of the final model. The overall ECD1 dataset was

divided into 64% (70916 images), 20% (22159 images), and 16% (17729 images) for training, testing, and validation, respectively. Also, the ECD was divided into 64% (8416 images), 20% (2630 images), and 16% (2104 images) for training, testing, and validation, respectively. Tables 2-3 show the ECD1 and ECD2 datasets, each disease with its total number of images in the training, validation, and test subsets

Step 4: The training phase of the ECDCNet model was applied in two approaches: The first approach utilizes CNNs without segmented lungs, while the second approach employs CNNs with segmented lungs (as depicted in Fig. 4). The utilization of both non-segmented and segmented lungs in the ECDCNet model ensures efficient detection of anomalies, including texture abnormalities, focal abnormalities, and abnormal shapes. In the first approach, the base CNNs were trained using image-based data without including segmented lungs. Conversely, the second approach trained CNNs using segmented-based images. Initially, the trained UNet model extracted the segmentation masks for the images to highlight the lungs. Subsequently, a dilatation function was applied to the segmented lung masks to address any potential breaks or intrusions. Additionally, the resulting dilated segmented lung mask was multiplied with the original image to extract the lung area to eliminate all background information. Ultimately, CNNs trained using the resulting images as input data

Step 5: Several methods were utilized to incorporate the base CNNs to derive the final prediction, including Majority Voting, Average Voting, and WPME

Table 2: The ECD1, each disease with its total number of images in the training, validation, and test subsets

Diseases	No. of images in the training/ validation/test subsets	Total
Cardiomegaly	1777/ 444/555	2776
Emphysema	1610/ 403/503	2516
Edema	1474/368/461	2303
Hernia	146/36/45	227
Pneumothorax	3394/848/1060	5302
Effusion	8523/2131/2663	13317
Mass	3701/925/1156	5782
Fibrosis	1079/270/337	1686
Atelectasis	7398/1849/2312	11559
Consolidation	2987/747/933	4667
Pleural thickening	2166/542, 677	3385
Nodule	4052/1013/1266	6331
Pneumonia	4763/1191/1489	7443
Infiltration	12732/3183/3979	19894
Covid-19	2314/579/723	3616
No finding	12800/3200/4000	20000
Total	70916/17729/22159	110804

Table 3: The ECD2, each disease with its total number of images in the training, validation, and test subsets

Diseases	No. of images in the training/validation/ test subsets	Total
Covid-19	2736/684/855	4285
Pneumonia	2480/620/775	3855
No finding	3200/800/1000	5000
Total	8416/2104/2630	13150

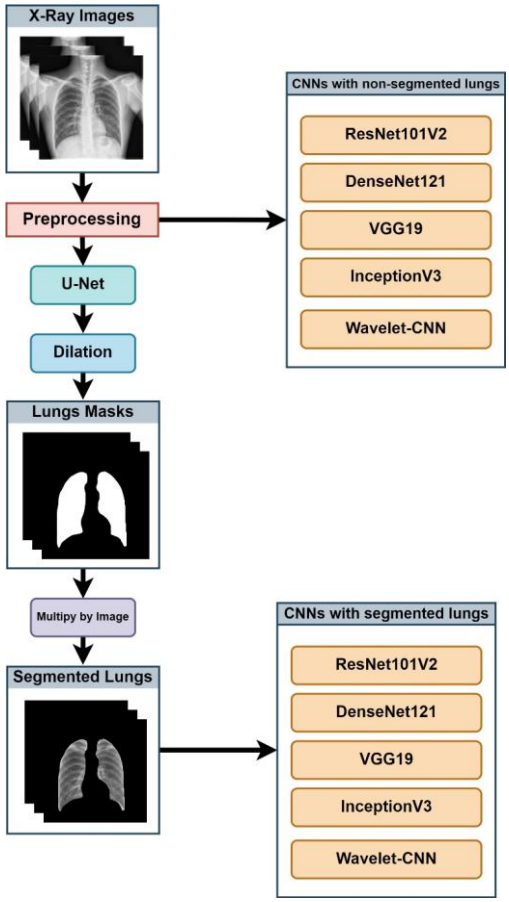

Fig. 4: The architecture of the proposed ECDCNet model in the training phase

The U-Net was trained by collecting a dataset of lung images and its ground truth masks that outline the lung regions, applying several processing techniques: Ensuring consistent resolution and reducing noise and splitting the dataset into training (64%), validation (16%) and testing subsets (20%). Then, Augment the training dataset to enhance the variability and the model's generalization ability: Rotation, scaling, and flipping. The model was trained for 100 epochs.

The Prediction Process

In the testing process, the ECDCNet model consists of two branches for evaluation. The first branch involves the trained CNNs without segmented lungs, while the

second branch involves the trained CNNs with segmented lungs. In the first branch, the image undergoes preprocessing and is inputted to all CNNs to get the prediction, as depicted in Fig. 5.

In the second branch, the image also undergoes preprocessing before being forwarded to the trained U-Net to obtain the lung segmentation mask. The resulting mask image is then dilated and multiplied by the original image to extract the lung area. Then, the processed image feds to all trained CNNs for prediction. The prediction results from the ten CNNs in both branches are combined using an ensemble unit. The ensemble unit utilizes the prediction results to generate a reliable final prediction result.

Performance Evaluation Metrics

The performance of the ECDCNet model and all its base CNNs were evaluated using various performance metrics. These metrics are Recall (R), Specificity (S), Accuracy (Acc), Precision (P), F1-score, Receiver Operating Characteristics (ROC) curve, and Area Under the Curve (AUC).

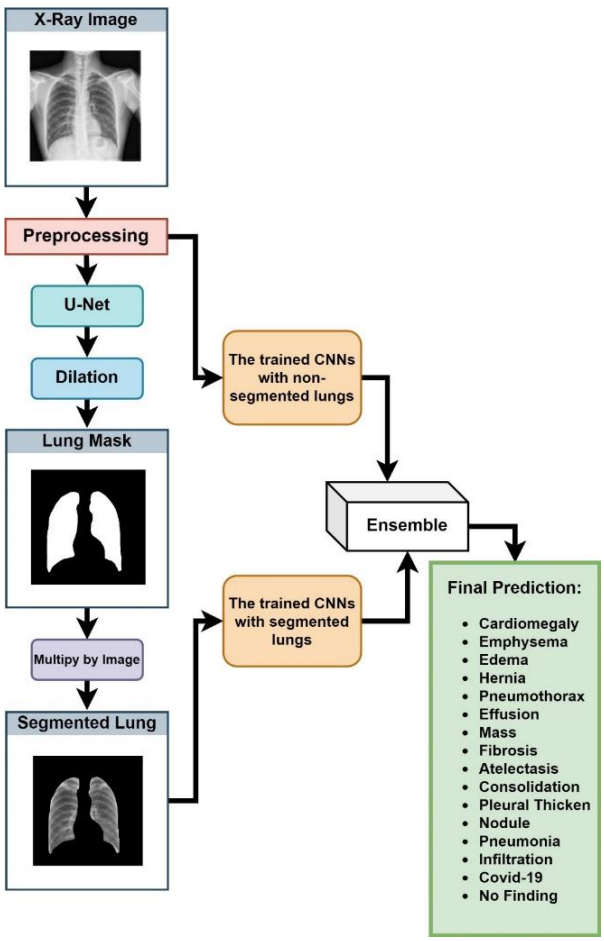

Fig. 5: The architecture of the proposed ECDCNet model in the prediction approach

Firstly, four terms will be explained which are True Positives (TP): The model will correctly predict the class of the image, True Negatives (TN): The model will correctly predict the images that do not belong to the class, False Positives (FP): The model will wrongly predict the class of the image and False Negatives (FN): The model will wrongly predict the images that are not belonging to the class.

Acc is the best way to know the overall performance of the model, which is calculated by finding the ratio between the correct predictions and the total predictions and it is calculated as in Eqs. 8-12 show how to calculate *R*, *S*, *P*, and *F1-scores*:

$$Acc = \frac{Numberofcorrectpredictions}{Totalofthesamples} \tag{8}$$

$$R = \frac{TP}{TP + FN} \tag{9}$$

$$S = \frac{TN}{TN + FP} \tag{10}$$

$$P = \frac{TP}{TP + FP} \tag{11}$$

$$F1 - score = \frac{2 * P * S}{P + S} \tag{12}$$

The *ROC* curve is a graph that shows the model's performance at various threshold levels. This curve is plotted between the True Positive Rate (TPR) and the False Positive Rate (FPR). Equations 13-14 show how to calculate the *TPR* and *FPR*, respectively. The AUC is the entire area under the ROC curve:

$$TPR = \frac{TP}{TP + FN} \tag{13}$$

$$FPR = \frac{FP}{FP + TN} \tag{14}$$

In image segmentation, the dice coefficient is doubled of the intersection area between two images divided by the sum of pixels in both images. Equation 15 shows how to calculate the dice coefficient:

$$Dice\ Cofficient = \frac{Area\ of\ intersection}{Total\ pixels\ in\ both\ images}$$

Results

This section presents the ECDCNet model's performance, base CNNs, U-Net, and ensemble strategies. The ECDCNet was implemented in Python using PyTorch, with software such as numpy, scipy, cv2, and pandas. Training included up to 50 epochs with early

stopping based on ROC. Hardware comprised a 32-Core/64-Thread Processor, 128GB DDR4 Memory (4x32GB), and a 960GB NVMe SSD.

Performance Results of Base CNNs

The ECDCNet model underwent training and evaluation using the collected datasets ECD1 and ECD2. All the ensemble approaches and the base CNNs of the ECDCNet model achieved the best performance metric with epochs less than 35 in ECD1 and less than 30 in ECD2.

ECD1 Dataset

Among all base CNNs, ResNet152V2 demonstrated the highest accuracy on the validation dataset. When using the segmented lung approach, it achieved an accuracy of 92.8%. In contrast, with the non-segmented lung approach, it achieved a slightly lower accuracy of 90.0%. On the other hand, InceptionV3 exhibited the lowest accuracy among the base CNNs on the validation dataset for ECD1. With the segmented lung approach, it achieved an accuracy of 84.9%, while with the non-segmented lung approach, it achieved an accuracy of 80.5%. Table 4 shows the results of the base CNNs with the non-segmented lung approach of ECD1. Table 5 shows the results of the base CNNs with the non-segmented lung approach in the validation dataset of ECD1.

In the test phase, the best-performing base CNN for the non-segmented lung approach was ResNet152V2. As for the segmented lung approach, VGG19 emerged as the top-performing base CNN.

Table 6 shows the results of the base CNNs with the non-segmented lung approach in the test dataset of ECD1. Table 7 shows the results of the base CNNs with the non-segmented lung approach in the test dataset of ECD1.

ECD2 Dataset

ResNet152V2 achieved the highest accuracy in the validation dataset among all base CNNs. It achieved an accuracy of 92.6% in the segmented lung approach and 91.9% in the non-segmented lung approach. Further, wavelet-CNN achieved the worst accuracy in the validation dataset in the non-segmented lungs approach. It achieved an accuracy of 88.9%. InceptionV3 achieved the worst accuracy in the validation dataset of ECD2 in the segmented lungs approach. It achieved an accuracy of 90.9%. Table 8 shows the results of the base CNNs with the non-segmented lung approach of ECD1. Table 9 shows the results of the base CNNs with the non-segmented lung approach in the validation dataset.

Table 4: Performance results of the base CNNs with non-segmented lungs approach in the validation dataset of ECD1

CNNs	Accuracy	F1-score	Precision	Specificity	Recall	AUC
VGG19	88.6	88.4	87.5	87.9	89.4	83.2
InceptionV3	80.5	88.3	89.5	80.8	87.2	82.0
DenseNet121	88.3	90.2	90.4	89.1	90.0	82.6
ResNet152V2	90.0	90.5	90.8	89.9	90.3	84.9
Wavelet-CNN	84.3	88.1	88.7	88.2	87.6	81.6

Table 5: Performance results of the base CNNs with segmented lungs approach in the validation dataset of ECD1

CNNs	Accuracy	F1-score	Precision	Specificity	Recall	AUC
VGG19	91.6	90.7	88.7	91.1	92.7	90.3
InceptionV3	84.9	91.2	91.4	90.7	91.1	87.9
DenseNet121	90.7	92.5	91.9	89.9	93.2	88.4
ResNet152V2	92.8	93.0	92.1	92.0	93.9	91.1
Wavelet-CNN	88.3	91.1	90.8	90.1	91.5	87.4

Table 6: Performance results of the base CNNs with non-segmented lungs approach in the test dataset of ECD1

	Accuracy	F1-score	Precision	Specificity	Recall	AUC
VGG19	89.5	89.4	88.3	89.9	90.5	84.2
InceptionV3	81.4	89.2	90.4	81.8	88.1	82.7
DenseNet121	89.7	91.0	91.4	90.3	90.7	83.5
ResNet152V2	90.8	91.4	91.6	91.3	91.2	86.2
Wavelet-CNN	86.4	89.1	89.6	89.2	88.7	82.9

Table 7: Performance results of the base CNNs with segmented lungs approach in the test dataset of ECD1

	Accuracy	F1-score	Precision	Specificity	Recall	AUC
VGG19	93.4	91.5	89.8	92.0	93.3	91.3
InceptionV3	86.7	91.9	92.1	91.4	91.8	88.5
DenseNet121	91.9	93.2	92.7	90.5	93.8	90.1
ResNet152V2	93.3	93.8	93.0	92.8	94.7	93.5
Wavelet-CNN	89.1	91.8	91.5	90.8	92.2	89.5

Table 8: Performance results of the base CNNs with non-segmented lungs approach in the validation dataset of ECD2

CNNs	Accuracy	F1-score	Precision	Specificity	Recall	AUC
VGG19	91.8	91.9	91.8	95.8	91.9	91.5
InceptionV3	89.1	88.9	89.1	94.6	88.9	90.8
DenseNet121	89.8	89.9	89.8	94.7	89.9	89.2
ResNet152V2	91.9	91.9	91.8	95.8	91.9	91.6
Wavelet-CNN	88.9	88.9	88.8	94.2	88.9	89.7

Table 9: Performance results of the base CNNs with segmented lungs approach in the validation dataset of ECD2

CNNs	Accuracy	F1-score	Precision	Specificity	Recall	AUC
VGG19	92.1	92.3	92.4	96.0	92.3	91.9
InceptionV3	90.9	91.0	91.1	95.3	91.0	91.4
DenseNet121	91.3	91.6	91.7	96.0	91.7	90.6
ResNet152V2	92.6	93.1	92.7	96.2	93.4	92.5
Wavelet-CNN	91.1	91.3	91.7	95.9	91.3	90.8

Table 10: Performance results of the base CNNs with non-segmented lungs approach in the test dataset of ECD2

	Accuracy	F1-score	Precision	Specificity	Recall	AUC
VGG19	94.9	95.0	95.1	97.0	94.8	95.8
InceptionV3	91.1	91.2	91.8	95.4	91.3	90.9
DenseNet121	92.5	92.4	92.4	96.1	92.5	92.3
ResNet152V2	93.5	93.5	93.4	96.7	93.7	90.5
Wavelet-CNN	93.3	93.4	93.4	96.5	93.4	90.3

Table 11: Performance results of the base CNNs with segmented lungs approach in the test dataset of ECD2

	Accuracy	F1-score	Precision	Specificity	Recall	AUC
VGG19	96.8	96.9	97.0	98.4	96.7	97.1
InceptionV3	95.3	95.0	95.1	97.6	94.9	95.8
DenseNet121	95.5	95.3	95.4	97.6	95.2	96.0
ResNet152V2	96.4	96.0	96.0	97.7	96.0	96.1
Wavelet-CNN	94.4	94.5	94.4	97.1	94.7	94.8

In the test dataset, the best base CNNs in the test dataset in both approaches, non-segmented and segmented lung, was VGG19. Table 10 shows the results of the base CNNs with the non-segmented lung approach in the test dataset of ECD1. Table 11 shows the results of the base CNNs with the non-segmented lung approach in the test dataset.

The ECDCNet Model Results

The ECDCNet Model accomplished excellent outcomes in all performance metrics with both the training and testing approaches. Figures 6-7 displayed accuracy training charts for the three ensemble models, indicating a positive trend in training accuracy over epochs. Moreover, it demonstrated accurate classification ability for all these ensembles.

All ensemble strategies predict more accurately than the individual CNNs in both datasets. The highest accuracy achieved in the test dataset belonged to the proposed ensemble strategy (WPME) in both ECD1 and ECD2. It achieved an accuracy of 96.1% in ECD1 and 98.9% in ECD2.

Tables 12-13 show the performance results for majority voting, average voting, and WPME in ECD1 and ECD2. Furthermore, in both ECD1 and ECD2, the WPME achieved the highest AUC values. It achieved an AUC of 99.5 in ECD2 and 98.3 in ECD1. His models exhibited a commendable execution time (time to test per image) of approximately less than 5 sec on a PC equipped with an 11th-Gen Intel Core i7 CPU and 32 GB DDR5 SDRAM.

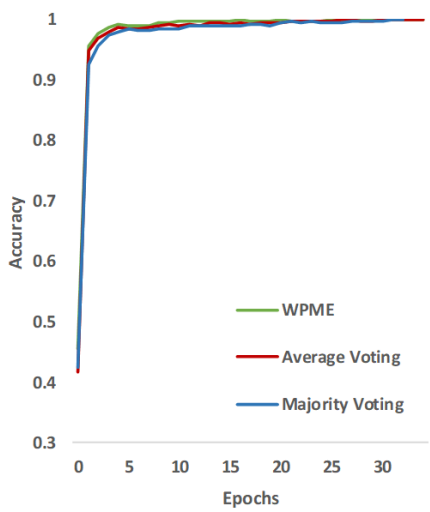

Fig. 6: The training accuracy graph for majority voting, average voting and WPME in the ECD1 dataset

Table 12: Performance results of the ECDCNet Model for majority voting, average voting, and WPME in the test dataset

	Accuracy	F1-score	Precision	Specificity	Recall	AUC
Majority voting	95.3	94.9	94.2	94.6	95.6	97.5
Average voting	95.8	96.0	95.3	95.8	96.8	97.9
WPME	96.1	96.5	95.8	96.3	97.2	98.2

Table 13: Performance results of the ECDCNet Model for majority voting, average voting, and WPME in the test dataset

	Accuracy	F1-score	Precision	Specificity	Recall	AUC
Majority voting	97.9	97.8	98.0	98.9	97.7	98.9
Average voting	98.2	98.2	98.0	99.0	98.3	99.1
WPME	98.9	98.9	98.7	99.4	99.0	99.5

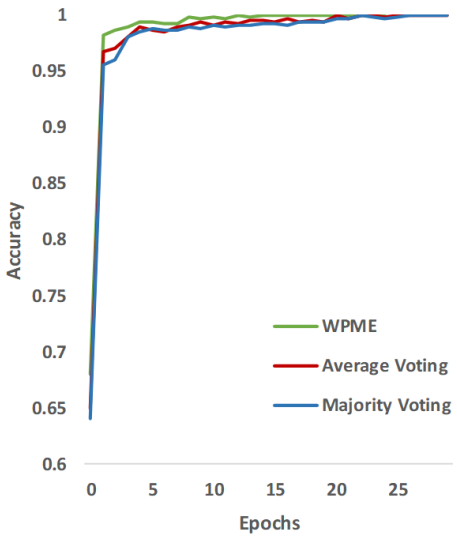

Fig. 7: The Training accuracy graph for majority voting, average voting, and WPME in the ECD2 dataset

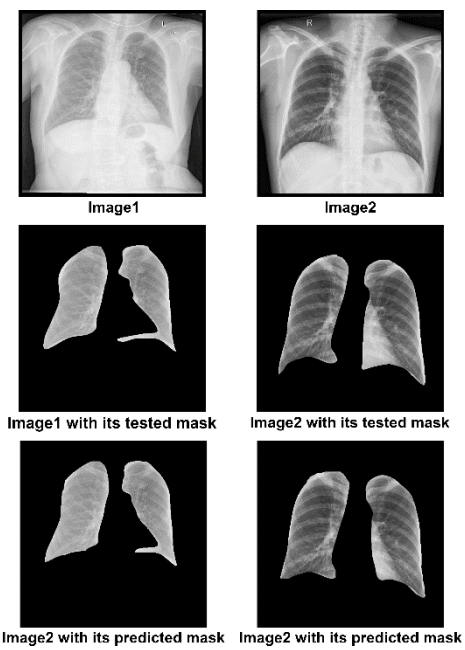

Fig. 8: Some of the segmented lungs produced by the U-Net

Performance Results of U-Net

The U-Net used in the ECDCNet model for lung segmentation was trained and evaluated using the "chest X-ray dataset for lung segmentation dataset", The AUC achieved by the U-Net was 99.1%, while accuracy, dice coefficient, specificity, recall, and precision were 96.2, 94.3, 98.3, 97.9 and 95.1%, respectively. The figure shows some of the segmented lungs that are produced by the U-Net. Figure 8 presents some of the segmented lungs that are produced by the U-Net.

Discussion

This study introduces a model comprising three CNN-ensemble strategies for classifying fifteen chest diseases using X-ray images. The model employed five CNN architectures as base classifiers based on their best performance results in recent studies. The chosen CNNs underwent two rounds of training: The first round involved standard X-ray images (non-segmented lungs), while the second round focused solely on the segmented lung area, extracted using a separate U-Net CNN.

To comprehensively evaluate the proposed model, various evaluation metrics were employed to assess the advantages and disadvantages of the ensemble strategies. Additionally, the performance of the base CNNs was assessed using both non-segmented X-ray images and segmented lung X-ray images.

From the obtained results, several valuable observations were made:

- The performance of disease classification was significantly improved by incorporating lung area segmentation, which allowed the convolutional neural networks (CNNs) to concentrate exclusively on the lung regions. Table 14 compares the accuracy of diagnosing chest diseases for the base CNNs with and without segmentation. The results demonstrate that all CNNs benefited from the lung segmentation technique, enhancing performance
- The ensemble strategies demonstrated their effectiveness by leveraging the combined performance of multiple CNNs, surpassing the individual CNN models in classification performance. In Table 14, it is

clear that all ensemble strategies predict more accurately than the individual CNNs

- Among the ensemble strategies and all the base CNNs, the WPME strategy demonstrated slightly superior performance achieving an accuracy of 96.1% in the sixteen-class classification and 98.9% in the three-class classification. The observed improvement can be attributed to utilizing four evaluation measures, precision, recall, F1-score, and accuracy, for setting the weights of base CNNs in the prediction stage, unlike the other strategies that solely relied on one measure
- The ECDCN model has consistently demonstrated outstanding performance in both datasets: ECD1 and ECD2. In the ECD2 dataset, which consisted of only three classes, our model achieved an impressive accuracy of 98.9%. Similarly, in the ECD1 dataset, encompassing 16 different classes, our model achieved a remarkable accuracy of 96.1%
- Finally, the overall results underscore the effectiveness of ensemble deep learning algorithms facilitating the early detection and diagnosis of chest diseases, which can aid in preventing disease progression

In order to assess the performance of the ECDCNet model in comparison to the state-of-the-art ensemble and individual CNN models, a comprehensive evaluation was conducted. Specifically, a comparative analysis was carried out between the ECDCNet model and existing studies that focused on three class classifications, namely COVID-19, pneumonia, and normal. Since there are currently no models in the state-of-the-art literature capable of diagnosing up to sixteen classes, the comparison was limited to these three classes. Table 15 displays a comparison of accuracy between the ECDCNet model proposed in this study and existing research categorized into individual models, ensemble models, and hybrid models.

The findings demonstrated that the ECDCNet model delivered outstanding results compared to the current studies. Its strengths lie in its ability to diagnose a larger number of diseases and its improved accuracy.

In this study, several significant contributions were made toward the classification of chest diseases using X-ray images:

- The study proposed a novel model of three CNN-ensemble strategies to classify fifteen chest diseases. The selection of five CNN architectures based on their previous performance ensured the inclusion of robust models in the proposed ensemble. This model achieved 96.1% in the sixteen-class classification and 98.9% in the three-class classification
- The study introduced the concept of training the selected CNNs twice, which proved crucial in improving classification accuracy. The first training involved non-segmented X-ray images, while the second training focused exclusively on the segmented lung area using a dedicated U-Net CNN
- Another significant contribution of this study lies in the proposed ensemble strategies using various evaluation metrics. To enhance the performance of lung disease diagnosing, ultimately benefiting patients and medical professionals in their diagnostic endeavors
- The study collected data from six different resources to create a comprehensive dataset of 130764 X-ray images, which play a crucial role in enhancing performance
- The study proposed model achieved high scores for all metrics accuracy, F1-score, precision, specificity, recall, and AUC. This remarkable efficiency renders the model exceptionally well-suited for deployment within a hospital environment, enabling the accurate diagnosis of chest diseases

Table 14: Comparison of accuracy value among the proposed ECDCNet model and its base CNNs

Model	No. of classes	Accuracy
VGG19	3	96.8*, 94.9
InceptionV3	3	94.1*, 91.1
DenseNet121	3	95.3*, 92.5
ResNet152V2	3	96.4*, 93.5
Wavelet-CNN	3	94.4*, 93.3
ECDCNet-majority voting	3	97.9%
ECDCNet-average voting	3	98.2%
ECDCNet-WPME	3	98.7%

Table 15: Comparison of accuracy value among a sample of existing studies with the proposed ECDCNet model

Individual models		
Study	Model	Accuracy %
Sharma *et al*. (2022a)	VGG19	91.67
	ResNet50	90.33
	MobileNetV2	76.00
	SqueezeNet	97.33
Pramanik *et al*. (2022)	DenseNet121	96.96
	ResNet101	88.01
	VGG-19	96.03
	InceptionV3	95.10
Kaye *et al*. (2021)	AlexNet	96.46
	GoogLeNet	96.20
	SqueezeNet	94.78
Ensemble Models Study		
Alshahrni *et al*. (2023)	CNN architectures	98.06
Goyal and Singh (2021)	CNN architectures	94.31
Mabrouk *et al*. (2022)	CNN architectures	93.91
Solaiman *et al*. (2022)	CNN architectures	96.25
Banerjee *et al*. (2022)	CNN architectures	94.55
Pramanik *et al*. (2022)	CNN architectures	98.47
Hybrid Models Study		
Aslan *et al*. (2021)	mAlexNet	98.70
Kedia and Katarya (2021)	CoVNet-19	98.28
Wu *et al*. (2021)	ULNe	95.35
Shaheed *et al*. (2023)	Hybrid-features and Random forest classifier	97.00
The proposed ECDCNet Model		
This study	Majority Voting	97.90
	Average voting	98.20
	WPME	98.90

Conclusion

This study proposed a novel ensemble convolutional neural networks model named ECDCNet to diagnose fifteen types of chest diseases by analyzing X-ray images of the lungs. The ECDCNet model encompasses five CNNs deployed using ensemble strategies. Furthermore, the study proposed a CNN-ensemble method called Weighted Performance Metrics Ensemble (WPME) to enhance the prediction capabilities of the ensemble model. This strategy incorporates four evaluation measures precision, recall, F1-score, and accuracy during the prediction stage.

This study found that the ensemble convolutional neural networks significantly enhanced the performance of chest disease diagnosing since the proposed ECDCNet achieved an accuracy of 96.1% for the sixteen-class classification and 98.9% for the three-class classification. The study owes its success to the successful proposed ensemble strategy that used the performance metrics of each CNN to elect the final prediction diseases. Furthermore, this study proved that training the CNNs using lung segmentation outperforms the training using the entire chest X-ray images since all the performance metrics were increased for each base CNN. Also, the achieved results recommend that the ensemble CNN models can be utilized to improve the efficiency and accuracy of the diagnosis process compared to the individual CNNs.

In addition, if considering the implementation of the ECDCNet model in hospitals, it is expected to assist radiologists in detecting and diagnosing chest diseases in the patient X-ray image (etc., pneumonia, atelectasis, cardiac hypertrophy, effusions, nodules, and COVID-19) in a fast, effective and accurate manner. Further, the early diagnosing stage of chest disease using the ECDCNet model can help prevent the chest disease from worsening.

Future Work

There are several future directions for research:

- Using transfer learning to enhance the performance of the CNN model
- Using the ECDCNet model for other diseases (etc. eye diseases, joint diseases, and phytopathology)
- Developing a real-time diagnosis system for chest diseases
- Exploring other CNNs architectures (etc. DenseNet201, InceptionResnetV2 and Faster R-CNN)
- Using different types of images for chest diseases (etc., CT and MRI)

Acknowledgment

We would like to thank all persons had contributed to this study. Also, we express our gratitude to the study's reviewers and editors.

Funding Information

This study did not receive funds from either public or private entities.

Author's Contributions

Each author contributed significantly to this study.

Mansour Ali Alhlalat: The study background and previous work, methodology, programming, implementation and discussion.

Ahmad Abdel-Aziz Sharieh: The review of the study methodology, implementation and discussion. Also, the study administration.

Mohammed Belal Al-Zoubi: The review of the study methodology, implementation and discussion.

Ethics

The corresponding author declared that this study has not been submitted elsewhere. Also, the authors declared no conflict of interest.

References

Agrawal, K., Kumar, R., & Jain, S. (2022). An efficient ensemble model for diagnosing COVID-19 and pneumonia using chest x-ray images. *Indian Journal of Science and Technology*, *15*(38), 1900-1906. https://doi.org/10.17485/IJST/v15i38.1269

Ait Nasser, A., & Akhloufi, M. A. (2023). A review of recent advances in deep learning models for chest disease detection using radiography. *Diagnostics*, *13*(1), 159. https://doi.org/10.3390/diagnostics13010159

Al Jumah, A. (2013). Denoising of an image using discrete stationary wavelet transform and various thresholding techniques. *Journal of Signal and Information Processing*, *4*(1), 33-41. https://doi.org/10.4236/jsip.2013.41004

Almezhghwi, K., Serte, S., & Al-Turjman, F. (2021). Convolutional neural networks for the classification of chest X-rays in the IoT era. *Multimedia Tools and Applications*, *80*(19), 29051-29065. https://doi.org/10.1007/s11042-021-10907-y

Alshahrni, M. M., Ahmad, M. A., Abdullah, M., Omer, N., & Aziz, M. (2023). An intelligent deep convolutional network based COVID-19 detection from chest X-rays. *Alexandria Engineering Journal*, *64*, 399-417. https://doi.org/10.1016/j.aej.2022.09.016

Arvind, S., Tembhurne, J. V., Diwan, T., & Sahare, P. (2023). Improvised light weight deep CNN based U-Net for the semantic segmentation of lungs from chest X-rays. *Results in Engineering*, *17*, 100929. https://doi.org/10.1016/j.rineng.2023.100929

Aslan, M. F., Unlersen, M. F., Sabanci, K., & Durdu, A. (2021). CNN-based transfer learning-BiLSTM network: A novel approach for COVID-19 infection detection. *Applied Soft Computing*, *98*, 106912. https://doi.org/10.1016/j.asoc.2020.106912

Banerjee, A., Sarkar, A., Roy, S., Singh, P. K., & Sarkar, R. (2022). COVID-19 chest X-ray detection through blending ensemble of CNN snapshots. *Biomedical Signal Processing and Control*, *78*, 104000. https://doi.org/10.1016/j.bspc.2022.104000

Bassi, P. R., & Attux, R. (2022). COVID-19 detection using chest X-rays: Is lung segmentation important for generalization?. *Research on Biomedical Engineering*, *38*(4), 1121-1139. https://doi.org/10.1007/s42600-022-00242-y

Bhatt, H., & Shah, M. (2023). A Convolutional Neural Network ensemble model for Pneumonia Detection using chest X-ray images. *Healthcare Analytics*, *3*, 100176. https://doi.org/10.1016/j.health.2023.100176

Bosdelekidis, V., & Ioakeimidis, N. S. (2020). Lung field segmentation in chest X-rays: A deformation-tolerant procedure based on the approximation of rib cage seed points. *Applied Sciences*, *10*(18), 6264. https://doi.org/10.3390/app10186264

Cao, F., & Zhao, H. (2021). Automatic lung segmentation algorithm on chest X-ray images based on fusion variational auto-encoder and three-terminal attention mechanism. *Symmetry*, *13*(5), 814. https://doi.org/10.3390/sym13050814

Chang, C. Y., Lin, T. K., Lin, C. W., & Cheng, H. T. (2022, July). Application of TransUNet for Segmenting Lung Mass from Chest X-ray Image. In *2022 IEEE International Conference on Consumer Electronics-Taiwan* (pp. 175-176). IEEE. https://10.1109/ICCE-Taiwan55306.2022.9869180

Chowdhury, M. E., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. B., ... & Islam, M. T. (2020). Can AI help in screening viral and COVID-19 pneumonia?. *IEEE Access*, *8*, 132665-132676. https://doi.org/10.1109/ACCESS.2020.3010287

Cohen, M., Levine, S. M., & Zar, H. J. (2022). World Lung Day: Impact of "the big 5 lung diseases" in the context of COVID-19. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, *323*(3), L338-L340. https://doi.org/10.1152/ajplung.00261.2022

Danilov, V., Proutski, A., Kirpich, A., Litmanovich, D., & Gankin, Y. (2022). Chest X-ray dataset for lung segmentation. *Mendeley Data*, V1. https://doi.org/10.17632/8gf9vpkhgy.1

Deb, S. D., & Jha, R. K. (2020, December). COVID-19 detection from chest X-ray images using ensemble of CNN models. In *2020 International Conference on Power, Instrumentation, Control and Computing (PICC)* (pp. 1-5). IEEE. https://doi.org/10.1109/PICC51425.2020.9362499

Dey, N., Zhang, Y. D., Rajinikanth, V., Pugalenthi, R., & Raja, N. S. M. (2021). Customized VGG19 architecture for pneumonia detection in chest X-rays. *Pattern Recognition Letters*, *143*, 67-74. https://doi.org/10.1016/j.patrec.2020.12.010

El Lel, T., Ahsan, M., & Haider, J. (2023). Detecting COVID-19 from Chest X-rays Using Convolutional Neural Network Ensembles. *Computers*, *12*, 105. https://doi.org/10.3390/computers12050105

Erdaw, Y., & Tachbele, E. (2021). Machine learning model applied on chest X-ray images enables automatic detection of COVID-19 cases with high accuracy. *International Journal of General Medicine*, 4923-4931. https://doi.org/10.2147/ijgm.s325609

Gadgil, S. U., Endo, M., Wen, E., Ng, A. Y., & Rajpurkar, P. (2021, August). CheXseg: Combining expert annotations with DNN-generated saliency maps for X-ray segmentation. In *Medical Imaging with Deep Learning* (pp. 190-204). PMLR. https://doi.org/10.48550/arXiv.2102.10484

Ganaie, M. A., Hu, M., Malik, A. K., Tanveer, M., & Suganthan, P. N. (2022). Ensemble deep learning: A review. *Engineering Applications of Artificial Intelligence*, *115*, 105151-105198. https://doi.org/10.1016/j.engappai.2022.105151

Goyal, S., & Singh, R. (2021). Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques. *Journal of Ambient Intelligence and Humanized Computing*, *14*, 3239-3259. https://doi.org/10.1007/s12652-021-03464-7

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770-778). https://doi.org/10.1109/CVPR.2016.90

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708). https://doi.org/10.1109/CVPR.2017.243

Jaszcz, A., Połap, D., & Damaševičius, R. (2022). Lung x-ray image segmentation using heuristic red fox optimization algorithm. *Scientific Programming*, *2022*, 1-8. https://doi.org/10.1155/2022/4494139

Kalaivani, S., & Seetharaman, K. (2022). A three-stage ensemble boosted convolutional neural network for classification and analysis of COVID-19 chest x-ray images. *International Journal of Cognitive Computing in Engineering*, *3*, 35-45. https://doi.org/10.1016/j.ijcce.2022.01.004

Kaye, A. D., Okeagu, C. N., Pham, A. D., Silva, R. A., Hurley, J. J., Arron, B. L., ... & Cornett, E. M. (2021). Economic impact of COVID-19 pandemic on healthcare facilities and systems: International perspectives. *Best Practice & Research Clinical Anaesthesiology*, *35*(3), 293-306. https://doi.org/10.1016/j.bpa.2020.11.009

Kedia, P., & Katarya, R. (2021). CoVNet-19: A Deep Learning model for the detection and analysis of COVID-19 patients. *Applied Soft Computing*, *104*, 107184. https://doi.org/10.1016/j.asoc.2021.107184

Kermany, D., Zhang, K., & Goldbaum, M. (2018). Labeled Optical Coherence Tomography (OCT) and chest x-ray images for classification. *Mendeley Data*, *2*(2), 651. https://doi.org/10.17632/rscbjbr9sj.2

Kim, Y. G., Kim, K., Wu, D., Ren, H., Tak, W. Y., Park, S. Y., ... & Li, Q. (2022). Deep learning-based four-region lung segmentation in chest radiography for COVID-19 diagnosis. *Diagnostics*, *12*(1), 101. https://doi.org/10.3390/diagnostics12010101

Kotei, E., & Thirunavukarasu, R. (2022, November). Ensemble technique coupled with deep transfer learning framework for automatic detection of tuberculosis from chest x-ray radiographs. In *Healthcare*, *10*(11), 2335, MDPI. https://doi.org/10.3390/healthcare10112335

Kumarasinghe, K. A. S. H., Kolonne, S. L., Fernando, K. C. M., & Meedeniya, D. (2022). U-Net Based Chest X-ray Segmentation with Ensemble Classification for Covid-19 and Pneumonia. *International Journal of Online & Biomedical Engineering*, *18*(7). https://doi.org/10.3991/ijoe.v18i07.30807

Kundu, R., Das, R., Geem, Z. W., Han, G. T., & Sarkar, R. (2021). Pneumonia detection in chest X-ray images using an ensemble of deep learning models. *PloS One*, *16*(9), e0256630. https://doi.org/10.1371/journal.pone.0256630

Lakhani, P., Mongan, J., Singhal, C., Zhou, Q., Andriole, K. P., Auffermann, W. F., ... & Shih, G. (2023). The 2021 SIIM-FISABIO-RSNA Machine Learning COVID-19 Challenge: Annotation and standard exam classification of COVID-19 chest radiographs. *Journal of Digital Imaging*, *36*(1), 365-372. https://doi.org/10.1007/s10278-022-00706-8

Lee, C. C., So, E. C., Saidy, L., & Wang, M. J. (2022). Lung Field Segmentation in Chest X-ray Images Using Superpixel Resizing and Encoder–Decoder Segmentation Networks. *Bioengineering*, *9*(8), 351. https://doi.org/10.3390/bioengineering9080351

Lee, Y. G., Lee, P. H., Choi, S. M., An, M. H., & Jang, A. S. (2021). Effects of air pollutants on airway diseases. *International Journal of Environmental Research and Public Health*, *18*(18), 9905. https://doi.org/10.3390/ijerph18189905

Mabrouk, A., Díaz Redondo, R. P., Dahou, A., Abd Elaziz, M., & Kayed, M. (2022). Pneumonia detection on chest X-ray images using ensemble of deep convolutional neural networks. *Applied Sciences*, *12*(13), 6448. https://doi.org/10.3390/app12136448

Malik, H., Anees, T., Din, M., & Naeem, A. (2023). CDC_Net: Multi-classification convolutional neural network model for detection of COVID-19, pneumothorax, pneumonia, lung Cancer and tuberculosis using chest X-rays. *Multimedia Tools and Applications*, *82*(9), 13855-13880. https://doi.org/10.1007/s11042-022-13843-7

Mann, M., Badoni, R. P., Soni, H., Al-Shehri, M., Kaushik, A. C., & Wei, D. Q. (2023). Utilization of Deep Convolutional Neural Networks for Accurate Chest X-Ray Diagnosis and Disease Detection. *Interdisciplinary Sciences: Computational Life Sciences*, 1-19. https://doi.org/10.1007/s12539-023-00562-2

Miller, L. A., Royer, C. M., Pinkerton, K. E., & Schelegle, E. S. (2017). Nonhuman primate models of respiratory disease: Past, present and future. *ILAR Journal*, *58*(2), 269-280. https://doi.org/10.1093/ilar/ilx030

Mohammed, A., & Kora, R. (2023). A comprehensive review on ensemble deep learning: Opportunities and challenges. *Journal of King Saud University-Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2023.01.014

Nillmani, Sharma, N. Saba, L. Khanna, N. N. Kalra, M. K. Fouda, M. M., & Suri, J. S. (2022). Segmentation-Based Classification Deep Learning Model Embedded with Explainable AI for COVID-19 Detection in Chest X-ray Scans. *Diagnostics*, *12*, 2132. https://doi.org/10.3390/diagnostics12092132

Parida, P., & Bhoi, N. (2017). Wavelet based transition region extraction for image segmentation. *Future Computing and Informatics Journal*, *2*(2), 65-78. https://doi.org/10.1016/j.fcij.2017.10.005

Prakash, J. A., Ravi, V., Sowmya, V., & Soman, K. P. (2023). Stacked ensemble learning based on deep convolutional neural networks for pediatric pneumonia diagnosis using chest X-ray images. *Neural Computing and Applications*, *35*(11), 8259-8279. https://doi.org/10.1007/s00521-022-08099-z

Pramanik, R., Dey, S., Malakar, S., Mirjalili, S., & Sarkar, R. (2022). TOPSIS aided ensemble of CNN models for screening COVID-19 in chest X-ray images. *Scientific Reports*, *12*(1), 15409. https://doi.org/10.1038/s41598-022-18463-7

Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S. B. A., ... & Chowdhury, M. E. (2021). Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Computers in Biology and Medicine*, *132*, 104319. https://doi.org/10.1016/j.compbiomed.2021.104319

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234-241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Sapijaszko, G., & Mikhael, W. B. (2018, August). An overview of recent convolutional neural network algorithms for image recognition. In *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)* (pp. 743-746). IEEE. https://doi.org/10.1109/MWSCAS.2018.8623911

Sarvamangala, D. R., & Kulkarni, R. V. (2022). Convolutional neural networks in medical image understanding: A survey. *Evolutionary Intelligence*, *15*(1), 1-22. https://doi.org/10.1007/s12065-020-00540-3

Schultheiss, M., Schmette, P., Bodden, J., Aichele, J., Müller-Leisse, C., Gassert, F. G., ... & Pfeiffer, D. (2021). Lung nodule detection in chest X-rays using synthetic ground-truth data comparing CNN-based diagnosis to human performance. *Scientific Reports*, *11*(1), 15857. https://doi.org/10.1038/s41598-021-94750-z

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

Sharma, C. M., Goyal, L., Chariar, V. M., & Sharma, N. (2022a). Lung disease classification in CXR images using hybrid inception-ResNet-v2 model and edge computing. *Journal of Healthcare Engineering*, *2022*. https://doi.org/10.1155/2022/9036457

Sharma, N., Saba, L., Khanna, N. N., Kalra, M. K., Fouda, M. M., & Suri, J. S. (2022b). Segmentation-based classification deep learning model embedded with explainable AI for COVID-19 detection in chest X-ray scans. *Diagnostics*, *12*(9), 2132. https://doi.org/10.3390/diagnostics12092132

Shih, G., Wu, C. C., Halabi, S. S., Kohli, M. D., Prevedello, L. M., Cook, T. S., ... & Stein, A. (2019). Augmenting the national institutes of health chest radiograph dataset with expert annotations of possible pneumonia. *Radiology: Artificial Intelligence*, *1*(1), e180041. https://doi.org/10.1148/ryai.2019180041

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv Preprint arXiv*: 1409.1556. https://doi.org/10.48550/arXiv.1409.1556

Singh, S., Kumar, M., Kumar, A., Verma, B. K., & Shitharth, S. (2023). Pneumonia detection with QCSA network on chest X-ray. *Scientific Reports*, *13*(1), 9025. https://doi.org/10.1038/s41598-023-35922-x

Sogancioglu, E., Murphy, K., Calli, E., Scholten, E. T., Schalekamp, S., & Van Ginneken, B. (2020). Cardiomegaly detection on chest radiographs: Segmentation versus classification. *IEEE Access*, *8*, 94631-94642. https://doi.org/10.1109/ACCESS.2020.2995567

Solaiman, I. A., Sanjana, T. I., Sobhan, S., Maria, T. S., & Rahman, M. K., (2022). X-Ray Classification to Detect COVID-19 using Ensemble Model. *14th International Conference on Agents and Artificial Intelligence (ICAART)*, 375-386. https://www.scitepress.org/Papers/2022/108472/108472.pdf

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-9). https://doi.org/10.1109/CVPR.2015.7298594

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2818-2826). https://doi.org/10.1109/CVPR.2016.308

Taye, M. M. (2023). Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. *Computation*, *11*(3), 52. https://doi.org/10.3390/computation11030052

Teixeira, L. O., Pereira, R. M., Bertolini, D., Oliveira, L. S., Nanni, L., Cavalcanti, G. D., & Costa, Y. M. (2021). Impact of lung segmentation on the diagnosis and explanation of COVID-19 in chest X-ray images. *Sensors*, *21*(21), 7116. https://doi.org/10.3390/s21217116

van Ginneken, B., Hogeweg, L., & Prokop, M. (2009). Computer-aided diagnosis in chest radiography: Beyond nodules. *European Journal of Radiology*, *72*(2), 226-230. https://doi.org/10.1016/j.ejrad.2009.05.061

Visuña, L., Yang, D., Garcia-Blas, J., & Carretero, J. (2022). Computer-aided diagnostic for classifying chest X-ray images using deep ensemble learning. *BMC Medical Imaging*, *22*(1), 178. https://doi.org/10.1186/s12880-022-00904-4

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 2097-2106). https://doi.org/10.1109/CVPR.2017.369

Win, K. Y., Maneerat, N., Sreng, S., & Hamamoto, K. (2021). Ensemble deep learning for the detection of Covid-19 in unbalanced chest X-ray dataset. *Applied Sciences*, *11*(22), 10528. https://doi.org/10.3390/app112210528

Wu, T., Tang, C., Xu, M., Hong, N., & Lei, Z. (2021). ULNet for the detection of coronavirus (COVID-19) from chest X-ray images. *Computers in Biology and Medicine*, *137*: 104834-104846. https://doi.org/10.1016/j.compbiomed.2021.104834

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging*, *9*, 611-629. https://doi.org/10.1007/s13244-018-0639-9

Zhang, S., Wen, L., Bian, X., Lei, Z., & Li, S. Z. (2018). Single-shot refinement neural network for object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4203-4212). https://doi.org/10.1109/cvpr.2018.00442

Zhao, H., Fang, Z., Ren, J., MacLellan, C., Xia, Y., Li, S., ... & Ren, K. (2022). SC2Net: A novel segmentation-based classification network for detection of COVID-19 in chest X-ray images. *IEEE Journal of Biomedical and Health Informatics*, *26*(8), 4032-4043. https://doi.org/10.1109/JBHI.2022.3177854