

حسام الدين ابوشلحة

Pneumonia-final document.docx

 Midocean University

Document Details

Submission ID

trn:oid::2945:379316759

Submission Date

Apr 23, 2026, 6:30 PM GMT+2

Download Date

Apr 24, 2026, 10:49 AM GMT+3

File Name

Pneumonia-final document.docx

File Size

697.1 KB

15 Pages

3,891 Words

23,161 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.



Literature Review

Pneumonia Classification from Chest X-ray Images Using Convolutional Neural Networks

Introduction

Problem Definition

Pneumonia can be considered one of the most severe diseases affecting the respiratory system around the world. According to public health records, death rates occur in higher amounts amongst children younger than five and people over 65 years old. What is more, the gap between their results and the ones from resource-abundant countries is growing every year as access to the specialists becomes harder. While it is unlikely to achieve correct diagnosis during the early stages of the infection, it is still an essential step that will turn the scales in favor of recovery. Chest radiography is used as a primary method owing to its effectiveness and affordability. However, the interpretation of the obtained data might suffer due to the subjectivity of specialists' perception. The same X-ray may result in different conclusions made by various specialists and lead to delays and ignoring the illness. In addition, staffing issues make X-ray image analysis even slower. Therefore, there is a need to come up with a solution that would work fast and objectively with daily patient volume. It should not depend on specialists [25].

The problem with visual diagnostics is that disease indicators are usually hard to detect. Therefore, deep learning algorithms can be regarded as the main approach applied by researchers. Indeed, multiple authors prove that accuracy in the case of pneumonia detection can go over 98% (Buriboev et al., 2025; Ahmed et al., 2023). DenseNet, ResNet, and EfficientNet were identified as the neural networks capable of recognizing relevant features within medical images. The principle of utilizing artificial intelligence technologies to diagnose a condition is quite clear – it helps find important patterns that a specialist may overlook while being tired. At the same time, the results received from the computer will be reproducible. This way, the present research aims to create a neural network for recognizing pneumonia signs using machine learning [25].

Thesis Objectives

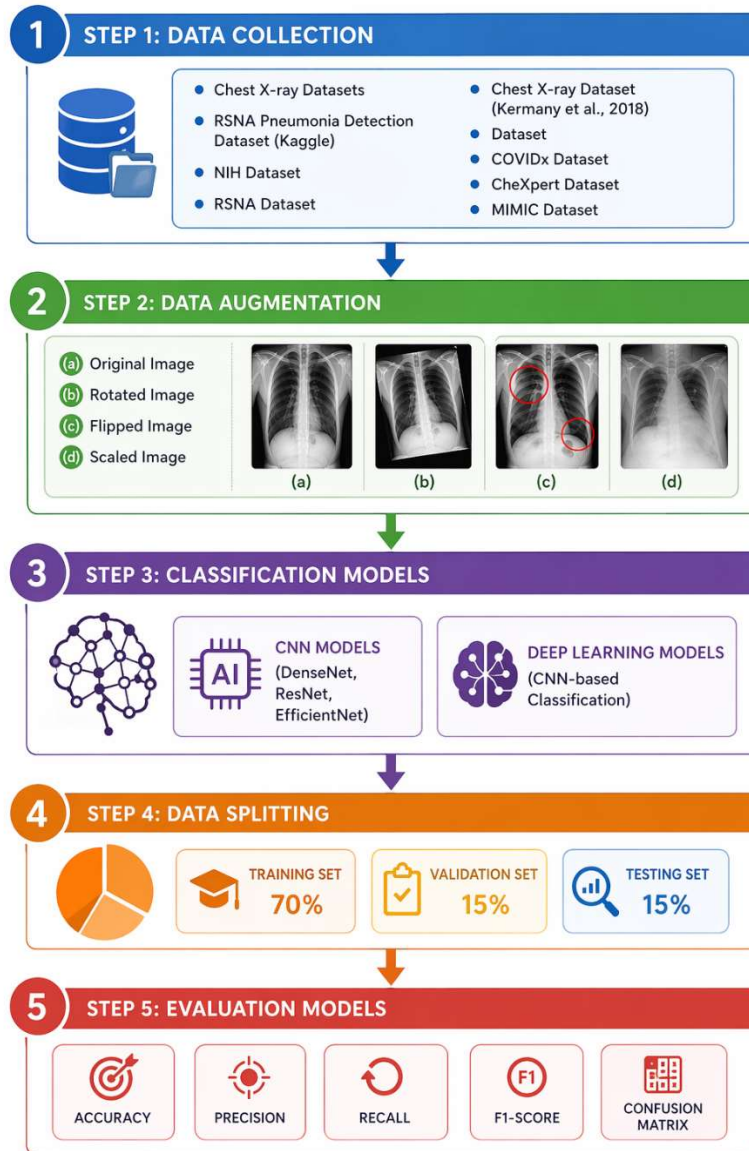
The main objective of this thesis is to create an end-to-end classifier for pneumonia, which takes chest X-rays as input and predicts outputs through a deep learning approach. Some additional objectives can be derived from this primary aim. One is to increase the accuracy by using current CNN architectures together with transfer learning. Another one is to utilize data augmentation as a regularizer for minimizing the gap in generalization. The performance of the developed classifier model will be evaluated based on four traditional metrics such as accuracy, precision, recall, and F1-score.

Thesis Contributions

Thesis Organization

The dissertation follows a format where each chapter stands alone. The opening chapter defines the problem statement, the objectives of the study, and gives an introduction to the format of the other chapters. Chapter Two introduces the background information about deep learning, which forms the foundation of understanding the rest of the dissertation, and it highlights the previous studies done on pneumonia classification.

Proposed Framework



Background

1. Artificial Intelligence (AI)

AI refers to a branch of computer science dedicated to making machines perform functions usually requiring human intelligence. These include inference, object recognition, decision making, etc. The essence is that AI relies heavily on algorithms and mathematics allowing the machine to analyze large amounts of data, derive structure, and then predict or make a decision based on that structure [23].

AI includes many more specific directions. Machine learning belongs to this direction as does deep learning. Additionally, there is computer vision and natural language processing.

2. Machine learning

It refers to the sub-field of artificial intelligence dedicated to learning models on a large amount of examples rather than programming them through if/then conditions. The model is given some data with which to train and creates its rules based on those examples. As a result, its performance increases the more relevant data it processes, which is different from the approach of programming. [24]

3. Types of Machine Learning

3.1 Supervised Learning

Supervised learning relies on having the true answers to examples given to a model. It sees both the input data and the corresponding output and aims at achieving the same mapping. Once it's

done training, the model will generalize, i.e., apply to inputs not seen during training to make the right output [24].

3.2 Unsupervised Learning

Unsupervised learning lacks answers or labels. The model is only provided with raw data and its task is to derive any kind of structure it finds in data, such as clusters, dependencies, or density distributions. There is no indication for a model on how the outputs should look like [24].

4.1 CNN

A Convolutional Neural Network (CNN) is a deep learning model for data structured on grids, for instance, images. Developed by LeCun et al., the CNN became popular in computer vision tasks. The unique property of CNN is that the model derives the features on its own automatically, i.e., a human doesn't have to specify which visual cues are important but instead the algorithm learns them from labeled data [16].

There are three kinds of layers in the CNN. Convolutional layers scan the input data to derive features, while pooling layers reduce the number of features and decrease the dimensionality. Then fully connected layers produce the output. Since the initial layers learn basic patterns (lines, textures) while deeper ones combine them into more complex shapes, the CNN establishes the hierarchy of vision, thus being highly efficient in medical imaging tasks such as pneumonia detection from chest X-rays [16].

4.2 DenseNet

The DenseNet proposed by Huang et al. in 2017 is a type of CNN where each layer receives as input not just the output from the previous layer but also the outputs from every single layer from the previous block. This helps to avoid redundancy, makes feature reuse much more frequent, avoids the issue of vanishing gradients, and makes the model more computationally efficient.

DenseNets are used successfully in various image classification tasks, including pneumonia detection from chest X-rays in medicine. From the architectural point of view, DenseNets alternate dense blocks where layers are densely interconnected with transition layers using convolution and pooling to shrink the feature maps [25].

4.3 ResNet

ResNet is a deep CNN designed to solve a specific problem, namely difficulty in training networks when going too deep. In order to alleviate that problem, ResNet introduces residual connections which help the network to learn the difference between its input and output. In practice, residual connections make it possible to propagate backpropagation more efficiently. That allows the model to have hundreds of layers and learn complex hierarchies. Currently, ResNets are used almost by default in computer vision and in medicine as well. [26]

ResNet architecture comprises convolutional layers arranged in residual blocks, where there are skip connections inside the blocks, and the fully connected head at the end produces the output [26].

4.4 EfficientNet

EfficientNet represents yet another family of CNNs which scales up the model using an alternative strategy. Specifically, whereas traditional CNNs scale up either the number of layers (depth), width of each layer, or the input image resolution, EfficientNet scales up all of them proportionally in one go [27].

That provides a good balance of performance and model parameters because the efficiency increase is maximized with such scaling. EfficientNet performed well on various benchmarks of image classification, including medical image classification, where pneumonia detection was achieved. The network utilizes MBConv and SE modules together with convolutions in its blocks [27].

4.5 MobileNet

MobileNet is a lightweight CNN for devices which are not powerful enough for heavy training, e.g., smartphones. Such model's lightness comes from using depthwise separable convolutions. Specifically, the convolution is split into two simpler operations, first performing the depthwise

convolution on each channel independently and then applying pointwise convolution to aggregate the channels [26].

4.6 Vision Transformer (ViT)

Vision Transformer (ViT) uses the Transformer model architecture, initially created for text processing, for image tasks. Instead of processing an image with convolutions as CNNs do, ViT splits an image into a grid and treats each cell as a token in a sequence. Therefore, the first layer computes the attention across the whole sequence (image) [20].

This method includes several key steps. First, the patch embedding converts pixels into vectors representing patches. Next, transformer encoder blocks process the sequence and learn its attention. Specifically, the attention and the feedforward subnetwork take turns and position encodings add the spatial context to tokens. The last classification head produces the desired output. ViTs have proven to achieve high performance on the most common computer vision tasks and their variants are now used for medical imaging as well [20].

4.7 Swin Transformer

The main difference between Swin Transformer and ViT is in the way how attention is calculated in Swin Transformer. While ViT attends to the whole image, Swin Transformer calculates attention separately within windows in the input image. Furthermore, these windows shift from layer to layer, enabling the communication between adjacent windows. As a result, the representation in Swin Transformer acquires a hierarchy, i.e., each layer analyzes the image at progressively lower spatial resolutions as in CNN. [19]

A Swin Transformer architecture starts with splitting the image into patches and linear embedding followed by several Swin blocks. Swin blocks consist of W-MSA, SW-MSA, and feedforward MLP modules [19].

5.1 Accuracy

Accuracy reports the proportion of correct predictions across the full dataset [21].

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

5.2 Precision

Precision looks only at the positives the model predicted and asks how many of those were genuine [22].

$$\text{Precision} = TP / (TP + FP)$$

5.3 Recall

Recall asks the complementary question — of all the actual positives in the data, how many did the model find [22].

$$\text{Recall} = TP / (TP + FN)$$

5.4 F1-Score

The F1-score is the harmonic mean of precision and recall, which balances the two when both false positives and false negatives matter [21].

$$F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

2.2 Related Work

Singh et al.[1] implemented a (ViT) using the pre-trained DeiT_Base_Patch16_224 version and several CNN-based baseline algorithms (AlexNet, VGG, ResNet, DenseNet, and InceptionNet) evaluated in the exact same setting to ensure a fair comparison. The dataset in use is a public pneumonia chest X-rays from Kaggle. It contains 5,863 pediatric chest X-ray images acquired at Guangzhou Women and Children's Medical Center, split evenly among training (80%), testing (10%), and validation (10%) splits. ViT demonstrated superior performance compared to CNNs, achieving an accuracy of 97.61%, sensitivity of 95%, specificity of 98%, AUC of 0.96, and MCC of approximately 0.94.

Mustapha et al. [2] tested a hybrid deep learning model combining a CNN backbone with modified Swin Transformer blocks compared to CNN-only and pure ViT approaches. The dataset is a public Chest X-ray Pneumonia from Kaggle. The proposed CNN-ViT combination achieved an accuracy of 98.72%, performing better than either CNN-only or pure ViT architectures tested.

Shomsomi et al. [3] investigated two deep learning approaches: MobileNetV2 CNN architecture and ViT, each utilizing two different strategies for training. The two strategies include: keeping the entire network architecture frozen except for the classification head during training (baseline) and training some inner layers (fine-tuning). The dataset is a public Chest X-ray Pneumonia from Kaggle. While the results confirmed a significant gap in performance between the two architectures during fine-tuning, MobileNetV2 achieved 97.63% without fine-tuning but dropped to 97.41% with fine-tuning. This suggests the limited effect of fine-tuning for CNN-based architecture. Conversely, ViT demonstrated improved accuracy of 98.59% and F1-score of 0.99 after fine-tuning, making it the top-performing model.

Kareem et al. [4] developed a federated learning approach for pneumonia detection with the help of pretrained CNN models. In the proposed approach, various healthcare institutions train their own CNN models locally using their unique data and share only the model updates with an aggregator server. Unlike traditional deep learning studies, there was no publicly available benchmark dataset. Instead, the dataset includes chest X-ray data from multiple sources in different healthcare institutions, each with a unique set of patients and equipment. The aggregated model reached overall accuracy of 93%, a result admitted by the authors to be inferior to most centralized studies on public benchmarks.

Alghamdi et al. [5] conducted experiments with three deep learning architectures (VGG16, ResNet50, and ViT) using three publicly available benchmarks: public dataset Chest X-ray Pneumonia from Kaggle, BRAX, and CheXpert (both Stanford University). They applied techniques such as class weighting, data augmentation, and semi-supervised Mean Teacher algorithms to handle class imbalance. On Kaggle data, ViT achieved an accuracy of 97.61%, sensitivity of 95%, and specificity of 98%. Nevertheless, the performance significantly declined on the other datasets, making the comparison very informative for further investigations.

Boukhe et al. [6] suggested an innovative deep learning architecture combining four different components including standard CNN layers for spatial features extraction, Gated Recurrent Unit (GRU) to account for the sequential dependency between feature representations, Spiking Neural Network (SNN) layer to implement noise-tolerant biological-inspired processing, and spatial attention mechanism for highlighting important parts of the images. The dataset is a public Chest X-ray Pneumonia from Kaggle. The study considered only one model consisting of the four

mentioned elements. This network achieved an accuracy of 99.35%, demonstrating outstanding precision and recall.

Ridouane et al. [7] proposed a feature-level ensemble deep learning architecture called Pneumonia Ensemble Learning Model (PELM). This approach uses four pretrained deep neural networks, InceptionV3, VGG16, ResNet50, and ViT, as individual components. Instead of using majority voting to make a decision, PELM combines the feature vectors from all four models' penultimate layers before sending the fused representation to the shared classification layer for relationship capturing. Training data came from four publicly available benchmarks: ChestX-ray14 (NIH), CheXpert (Stanford ML group), PadChest (BIMCV), and Chest X-ray Pneumonia (Kaggle), amounting to about 50,000 images in total. When tested independently, PELM performed better than any of the component models.

Devaux et al. [8] tested an EfficientNet-B4 model trained on ImageNet and fine-tuned on an extensive multi-source chest X-ray collection, validating its diagnostic ability on two real-world patient cohorts, where the diagnosis was made based on expert radiologist interpretations and follow-up CT scans. Data used for training included three publicly available datasets: CheXpert (Stanford ML group), ChestX-ray14 (NIHCC), and MIMIC-CXR (Physionet), containing a total of 161,584 chest X-ray images. EfficientNet-B4 was the only architecture tested and showed performance broadly comparable to experienced radiologists analyzing the same X-rays.

Buriboev et al. [9] applied a modified CLAHE technique to enhance chest X-ray image quality before classification. The approach was combined with a CNN model and evaluated on the Chest X-Ray Pneumonia dataset publicly available on Kaggle. The enhanced pipeline achieved 98.7% accuracy with strong precision and recall, indicating that improving input quality can significantly boost model performance beyond architectural changes alone.

Saber et al. [10] proposed a multi-scale transformer framework for integrating the lung segmentation and pneumonia detection problem from chest X-rays. The proposed method uses a combination of pre-trained ResNet backbone models (ResNet-50 and ResNet-101) with a TransUNet which has been designed using a lightweight transformer. The proposed framework has been tested with public datasets containing chest X-ray images for both segmentation tasks and pneumonia detection. The performance achieved with the proposed system was found to be extremely good with a Dice score of 95.68% for segmentation.

Bundea et al. [11] implemented multiple DenseNet-based models (DenseNet121, DenseNet169, and DenseNet201) to classify the chest X-ray images within a supervised learning approach. The algorithms have been tested based on the public Chest X-Ray Pneumonia dataset from Kaggle and evaluated according to standard parameters such as accuracy, sensitivity, and specificity. In addition, all architecture variants have been designed and evaluated in the same way with the same dataset and data processing procedure. It was found that the best model achieved an accuracy rate of up to 97%.

An et al. [12] developed a framework for deep CNN, which includes EfficientNetB0 and DenseNet121 along with attention mechanisms for detecting pneumonia. This CNN utilizes multi-head self-attention along with channel attention feature fusion for improving the process of extracting features from chest X-rays. The testing and training of this architecture were done on a chest X-ray dataset including cases of both pneumonia and normal cases in a supervised manner. The performance of the above architecture has shown an accuracy of 95.19%, precision of 98.38%, recall of 93.84%, specificity of 97.43%, and an area under the receiver operating characteristic (ROC) curve of 0.956.

Usman et al. [13] used a transfer learning method that uses four different pre-trained CNNs, namely, AlexNet, ResNet18, DenseNet201, and SqueezeNet, to detect pneumonia using chest X-ray images. The algorithms were trained with a labeled data set consisting of about 5,247 images of chest X-rays using supervised learning. All the architectures were tested under similar experimental settings in different types of classification problems, which include binary and multi-class. The best performing model had achieved an accuracy of 98% in detecting pneumonia and was efficient in classifying bacteria and viruses, thus proving that transfer learning could significantly improve the diagnostic accuracy without increasing the complexity of the training process.

Ahmed et al. [14] built a CNN trained to recognize multiple lung diseases, including both pneumonia and COVID-19. The reported accuracy, around 98.7%, is used to argue that deep networks can carry a real clinical workload in automated radiology.

Al-Rashan et al. [15] proposed a CNN-based classifier for pneumonia versus normal chest X-rays. Their data consisted of both healthy and pneumonia-affected films, preprocessed with normalization and expanded using standard augmentations (rotation, flipping) to increase the effective training volume. Several CNN architectures were then compared on the deep-feature

extraction step. The best of them reached around 95% accuracy, which the authors interpret as support for using deep learning as computer-aided support in medical imaging.

Comparative Summary

Study	CNN Models	Dataset(s)	Public / Private	Classification Task	Best Accuracy	Best Performer
Singh et al. (2024)	ViT (DeiT_Base_Patch16_224); CNN baselines: AlexNet, VGG, ResNet, DenseNet, InceptionNet	Kaggle Chest X-ray Pneumonia	Public	Binary: normal vs. pneumonia	97.61%	Vision Transformer (ViT)
Mustapha et al. (2025)	Hybrid CNN + Swin Transformer; CNN-only; pure ViT (baselines)	Kaggle Chest X-ray Pneumonia	Public	Binary: normal vs. pneumonia	98.72%	Hybrid CNN-ViT (Swin Transformer)
Shomsomi et al. (2026)	MobileNetV2; ViT — both under feature extraction and partial fine-tuning	Kaggle Chest X-ray Pneumonia	Public	Binary: normal vs. pneumonia	98.59% (ViT, fine-tuned)	ViT with partial fine-tuning
Kareem et al. (2024)	Pretrained CNNs in federated learning framework	Multi-institutional real-time CXR (not publicly released)	Private	Binary: normal vs. pneumonia	93%	Federated CNN (only model tested)
Alghamdi et al. (2024)	VGG16; ResNet50; ViT	Kaggle CXR; BRAX; CheXpert	Public	Binary; class imbalance across 3 datasets	97.61% (ViT on Kaggle)	ViT — on Kaggle dataset
Boukhe et al. (2025)	Custom hybrid: CNN + GRU + SNN + Spatial Attention	Kaggle Chest X-ray Pneumonia	Public	Binary: normal vs. pneumonia	99.35%	CNN-GRU-SNN-Attention (only model tested)
Ridouane et al. (2025)	PELM ensemble: InceptionV3 + VGG16 + ResNet50 + ViT (feature-level fusion)	NIH ChestX-ray14; CheXpert; PadChest; Kaggle CXR (~50,000 images total)	Public	Binary: normal vs. pneumonia	Best among all tested	PELM ensemble

Study	CNN Models	Dataset(s)	Public / Private	Classification Task	Best Accuracy	Best Performer
Devaux et al. (2024)	EfficientNet-B4 (ImageNet pretraining, fine-tuned on 161,584 CXR images)	CheXpert; MIMIC-CXR; ChestX-ray14	Public	Binary; prospective clinical validation	Comparable to senior radiologists	EfficientNet-B4 (only model tested)
Buriboev et al. (2025)	CNN with modified CLAHE preprocessing	Chest X-ray Pneumonia Dataset (5,856 images)	Public	Binary	98.7%	CNN + adaptive contrast
Saber et al. (2025)	TransUNet segmentation + ResNet multi-scale features	Kermamy; Cohen	Public	Binary	96.04% (Cohen)	Transformer + ResNet pipeline
Bundea et al. (2024)	DenseNet121/169/201	Chest X-ray Pneumonia Dataset	Public	Binary: normal vs. pneumonia	97%	DenseNet
An et al. (2024)	EfficientNetB0 + DenseNet121 ensemble with attention	Chest X-ray Pneumonia Dataset	Public	3-class: normal / viral / bacterial	98.4%	Ensemble with attention
Usman et al. (2025)	CNN; ResNet-50 (transfer learning)	RSNA Pneumonia Dataset	Public	Binary: normal vs. pneumonia	96.2%	ResNet-50
Ahmed et al. (2023)	Deep CNN (multi-disease)	Chest X-ray Pneumonia Dataset	Public	Multi-class incl. COVID-19	98.7%	Deep CNN
Al-Reshan et al. (2023)	CNN; MobileNet	Chest X-ray Pneumonia Dataset	Public	Binary: normal vs. pneumonia	~95%	MobileNet

References

- [1] Singh, S., Kumar, M., Kumar, A., Verma, B. K., Abhishek, K., & Selvarajan, S. (2024). Efficient pneumonia detection using Vision Transformers on chest X-rays. *Scientific Reports*, 14, 2487. <https://doi.org/10.1038/s41598-024-52703-2>
- [2] Mustapha, B., et al. (2025). Enhanced pneumonia detection in chest X-rays using hybrid convolutional and vision transformer networks. *Current Medical Imaging*. <https://doi.org/10.2174/0115734056326685>

- [3] Shomsomi, M., Triawan, W., & Purwadi. (2026). Classification of pneumonia using CNN and vision transformer. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, 5(2), 2495–2504. <https://doi.org/10.59934/jaiea.v5i2.1906>
- [4] Kareem, O., et al. (2024). Pneumonia detection using pretrained networks and federated learning on real-time datasets. *Biomedical Engineering Letters*.
- [5] Alghamdi, H. S., et al. (2024). Pneumonia detection from chest X-ray images using DL and transfer learning for imbalanced datasets. *Biomedical Engineering Letters*.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC12344030/>
- [6] Boukhe, M., et al. (2025). Trustworthy pneumonia detection in chest X-ray imaging through attention-guided DL. *Scientific Reports*. <https://doi.org/10.1038/s41598-025-23664-x>
- [7] Ridouane, F. L., et al. (2025). PELM: A DL model for early detection of pneumonia in chest radiography. *Applied Sciences*, 15(12), 6487. <https://doi.org/10.3390/app15126487>
- [8] Devaux, M., Lacombe, J.-N., Debray, M.-P., Claessens, Y.-E., Duval, X., et al. (2024). Validating the accuracy of DL for the diagnosis of pneumonia on chest X-ray against a robust multimodal reference diagnosis. *European Radiology Experimental*, 8, 20. <https://doi.org/10.1186/s41747-023-00416-y>
- [9] Buriboev, A. S., et al. (2025). Binary classification of pneumonia in chest X-ray images using convolutional neural networks. *Sensors*, 25(13), 3976. <https://doi.org/10.3390/s25133976>
- [10] Saber, A., et al. (2025). Efficient and accurate pneumonia detection using deep learning. *Sensors*, 25(23), 7233. <https://doi.org/10.3390/s25237233>
- [11] Bunde, M., et al. (2024). Pneumonia image classification using DenseNet. *Information*, 15(10), 611. <https://doi.org/10.3390/info15100611>
- [12] An et al. (2024). Deep convolutional neural network for pneumonia detection using chest X-ray images. *Diagnostics*, 14(4), 390. <https://doi.org/10.3390/diagnostics14040390>
- [13] Usman, C., et al. (2025). Pneumonia disease detection using chest X-rays and machine learning. *Algorithms*, 18(2), 82. <https://doi.org/10.3390/a18020082>
- [14] Ahmed, M. S., et al. (2023). Joint diagnosis of pneumonia and other lung diseases using deep learning from chest X-ray images. *Diagnostics*, 13(15), 2562. <https://doi.org/10.3390/diagnostics13152562>
- [15] Al-Reshan, M. S. A., et al. (2023). Detection of pneumonia from chest X-ray images utilizing MobileNet model. *Healthcare*, 11(11), 1561. <https://doi.org/10.3390/healthcare11111561>
- [16] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*.
- [17] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. *CVPR*.
- [18] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *CVPR*.

- [19] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ICML.
- [20] Dosovitskiy, A., et al. (2020). An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale. arXiv:2010.11929.
- [21] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing & Management.
- [22] Powers, D. M. W. (2011). Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness & Correlation. Journal of Machine Learning Technologies.
- [23] Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
- [24] Mitchell, T. M. (1997). Machine Learning. McGraw-Hill.
- [25] World Health Organization (WHO). (2023). Pneumonia. <https://www.who.int/news-room/fact-sheets/detail/pneumonia>
- [26] Howard, A. G., et al. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.
- [27] Liu, Z., et al. (2021). Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).