https://doi.org/10.48047/AFJBS.7.5.2025.913-928

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

Lightweight Hierarchical Attention-Compression: An Intelligent Deep Learning Approach for Disease Classification in X-ray Images

MUNESH MEENA, *, Ruchi Sehrawat[2]

University School of Information Communication & Technology, GGSIPU, Delhi, India

munesh8204@gmail.com[1,*], ruchi.sehrawat@ipu.ac.in[2]

Volume 7, Issue 5, May 2025

Received: 15 Mar 2025

Accepted: 05 Apr 2025

Published: 09 May2025

doi:10.48047/AFJBS.7.5.2025.913-928

Abstract:

Objective: Clinical radiography needs automated X-ray disease classification, especially in resource on strained settings. The primary aim is to improve diagnosis accuracy and reduce the manual burden on medical professionals. Existing methods have trouble combining information from different scales of features while simultaneously lowering the computational complexity. This makes it difficult to capture complex spatial and channel-wise correlations. Method: A novel and extremely lightweight HACNet model with just a million parameters has been
proposed to overcome these limitations. It contains a novel Hierarchical Attention-Compression Module (HACM) to boost feature representational capacity. A Feature Aggregation Unit (FAU)combines different feature maps using close connections and a Triplet Attention layer, while the Attentional Compression Unit (ACU) simplifies the data by using stride convolution along with channel and spatial attention methods.

Results: Results of experiments concerning the Covid-19 radiography dataset and the ChestX-ray8dataset demonstrate strong classification capabilities of the suggested model in identifying diseases from x-ray images. The suggested model is more computationally efficient than popular computer
vision models and demonstrates excellent qualitative capabilities, as demonstrated by t-SNE plots.

Keyword: x-ray, disease, classification, deep learning, attention, light-weight

Introduction Diagnostic radiology has mostly depended on quick, cost-effective and non-invasive X-ray imaging to view internal body structures. Clinics commonly use it to detect conditions in the chest area, including pneumonia [1]. TB, lung cancer, and viral infections like COVID-19 [2]. X-rays identify fractures, bone wear and tear, unusual bone formations, an enlarged heart, and fluid buildup in the lungs outside the chest. This shows the need for better methods of detecting issues using X-rays. In the early stages of pneumonia and TB, which are two of the most widespread infections in the world, the signs on X-rays are often very subtle and simple to miss. Signs like solid white patches, hollowed areas or fine patterns on X-rays can point to these diseases. Apart from infections, X-rays are also very important in spotting cancers such as lung cancer, which can greatly influence how treatment is planned. Reading the images correctly is key for quick diagnosis, proper patient care and effective treatment.

Because standard X-ray reading methods have their limits, automated detection methods [3] have become essential [4]. Differences in how well radiologists perform, tiredness and too many patients can lead to errors and delays in diagnosis, especially in places with limited resources. Automated systems that use strong algorithms can make the readings more reliable, reducing mistakes and ensuring every patient receives the same level of care. Adding these tools to regular medical practice can provide an extra opinion which is helpful when a trained radiologist is not available. Thanks to improvements in imaging and digital X-ray technology, high-quality images can now reveal even the tiniest details in body structures.

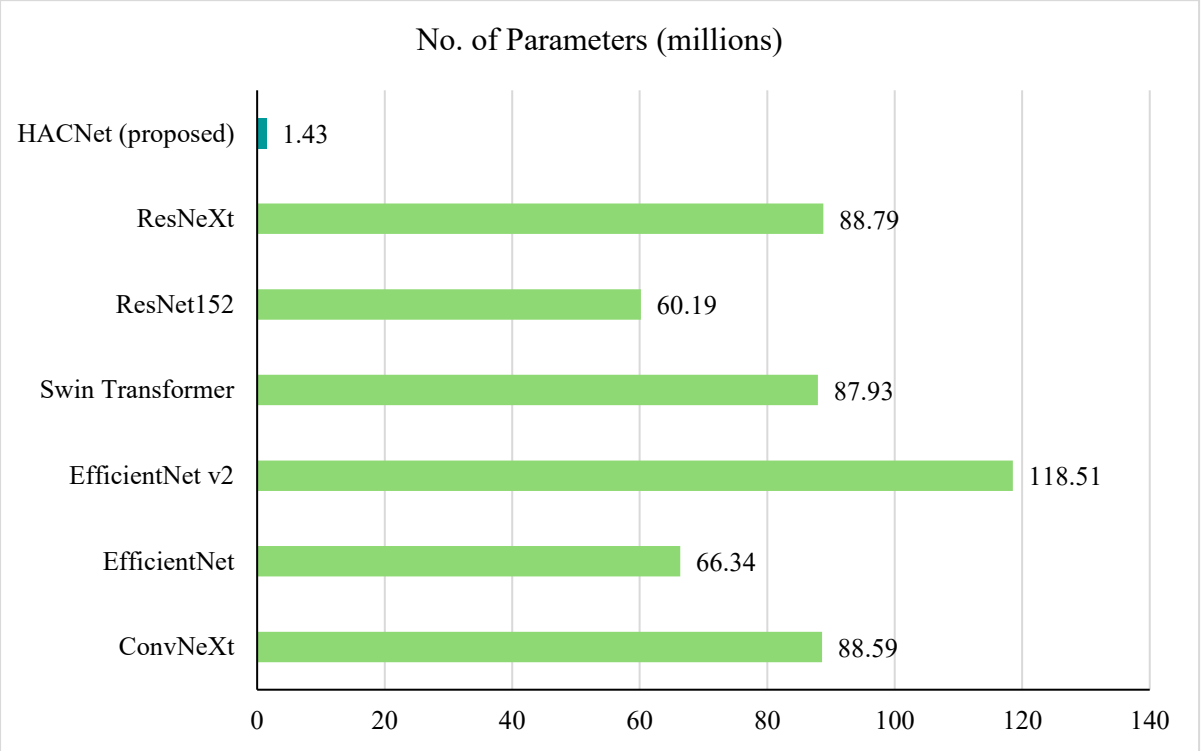


Fig. 1 Comparison of the number of papermeters in the proposed HACNet model against other popular computer vision models. The proposed model is extremely lightweight with only 1.43 million parameters, making it easy to train and deploy for disease detection.

Because the datasets are huge and complex manual analysis is difficult and effective computer-based methods for processing and analysing these images are essential. By using automated detection techniques, clinicians can rapidly examine large datasets, give priority to

urgent cases, and improve healthcare's overall delivery. As medical imaging and diagnostic technologies continue to advance, there is a growing need for strong and scalable detection systems. Automated X-ray analysis can reduce the workload for healthcare professionals, improve diagnostic accuracy, and make radiology departments more efficient [5]. These tools can support remote diagnostics and telemedicine to make quality healthcare accessible in underserved regions and such developments will improve patient outcomes lower medical costs and promote healthcare fairness.

Deep learning, and in particular convolutional neural networks, have transformed medical image analysis by automatically learning complex features from raw imaging data [6]. DenseNet, ResNet, and Inception are recently developed frameworks that feature deeper networks and more intricate layer connections. These architectural advances enable them to learn the structural patterns present in X-ray images. Transfer learning leverages models that were pretrained on large datasets such as ImageNet. These models are then fine-tuned with medical data to adapt them to the target task. This approach helps overcome the limitations of manually labelling medical datasets.

In deep learning models, feature extraction employs convolutional layers along with pooling techniques and nonlinear activation functions to detect variations in image texture [7] and spatial structure. These models are carefully designed to detect subtle changes in radiographic patterns that may signal disease conditions, thereby improving diagnostic accuracy. Convolutional neural networks can automatically learn important features, which marks a significant improvement over traditional image processing method.

Modern deep learning models now include attention mechanisms [8] to make X-ray interpretation more intuitive and to help the models focus on clinically significant regions. Attention modules enable the network to prioritise specific areas in the image and highlight features that are critical for diagnosis and recent improvements employ transformer-based designs and hybrid models that fuse convolutional networks with recurrent neural networks or graph techniques to capture relationships in data across space and time. These innovations contribute to a deeper understanding of imaging biomarkers associated with disease.

Current developments in medical imaging deep learning emphasise understandable AI and semi-supervised learning strategies, with GradCAM saliency maps and layer-wise relevance propagation showing how these models reach their decisions. This transparency builds trust and confidence among doctors. Integrating X-ray images with clinical information through multimodal data fusion and using deep learning algorithms that refine automated disease classification from radiographs provide clinicians with a comprehensive patient view and enable faster and more accurate diagnostics that enhance care.

This paper addresses these gaps by proposing a new deep learning model called HACNet. The model employs Hierarchical Attention-Compression Modules to improve feature fusion and compression. We improved the DenseNet design by adding attention mechanisms in both the denseblock and transition layers, which ultimately helps us achieve an optimal balance between speed and accuracy in diagnosis. The model undergoes extensive evaluations on standard datasets and proves capable of effective X-ray disease categorisation.

The key contributions of this paper are:

1. A lightweight HACNet deep learning model is proposed for disease classification from X-ray images.
2. Introduces HACM that synergistically combines dense feature aggregation and attention-driven compression, enhancing the ability of the model to capture robust, multi-scale spatial and channel-wise features from input images.
3. It connects convolutional layers closely and add a triplet attention layer that helps reuse features and creates a clear representation by understanding complex relationships at different levels.
4. Computation comparison proves that HACNet is extremely lightweight with just 1.43 million parameters and is easy to train and deploy.
5. The experimental results from the Covid19 Radiography and ChestX-ray8 datasets show that the HACNet model works well for classifying diseases. Qualitative analysis shows the classification of the model visually.

1 Literature Review

The review of the literature on disease classification from X-ray images is presented in this part.

Mei et al. [9] introduce HC-Net, a hybrid classification framework that accurately classifies periodontal disease from panoramic X-ray images using tooth-level and patient-level analysis. The framework applies instance segmentation to detect and classify each tooth, while a multitask approach learns global classification and generates a Classification Activation Map to identify lesion regions at the patient level. An adaptive noisy-OR gate combines predictions from both levels to address inconsistencies in manual annotation and enhance the sensitivity of periodontal disease diagnosis.

Ozturk et al. [10] propose HydraViT, which integrates a transformer backbone with a multibranch output module that uses learnt weighting to improve multilabel classification. The transformer component uses self-attention to capture long range contextual dependencies in X ray images and focuses on regions of diagnostic importance. The multibranch output module includes individual branches for each disease label and an aggregated branch to preserve sensitivity to co-occurring pathologies.

Prasath et al. [11] present DTRSN-XRI-CPI, an enhanced double transformer residual super-resolution network for pneumonia classification using chest X-rays. The model begins by preprocessing the images through region-aware neural graph collaborative filtering, which reduces noise, enhances contrast and removes irrelevant frequency components. Then the synchro squeezed fractional wavelet transform is applied to extract colour shape, spatial texture and relational features. Finally, hunter-prey optimisation algorithms are used to adjust the model weight parameters and enhance classification performance.

Zhao et al. [12] present an improved double transformer residual super-resolution network for pneumonia classification from chest X-ray. The synchro squeezed fractional wavelet transform extracts colour shape, spatial texture and relational features from images that have been pre-processed by region-aware neural graph collaborative filtering to reduce noise, enhance contrast and filter out unwanted frequency components. Hunter prey optimisation algorithms adjust the model weight parameters to boost classification performance.

Patankar et al. [13] enhance feature extraction for classification tasks through effective image preprocessing and use a spiking neural network trained with the tempotron learning rule to maintain robust temporal dynamics. The training process also incorporates standard time-dependent plasticity to improve processing speed and reduce complexity compared to traditional spiking network methods.

Rabbah et al. [14] develop a new CNN architecture trained on a curated dataset of 5856 chest X-ray images and use an Inception v3 layer for feature extraction followed by dense layers for binary classification to automate pneumonia diagnosis. To reveal its decision-making process, the model uses integrated gradients to highlight image regions that most influence its predictions. With 22.9 million trainable parameters, this approach achieves 97.23% diagnostic accuracy while enhancing clinical transparency.

Yi et al. [15] propose a BSD framework for bone suppression pulmonary parenchyma extraction and disease detection. This framework reduces the impact of bone structures and other tissues on lung disease diagnosis. IU-Net, a novel segmentation network, extracts bone-free pulmonary parenchyma by replacing standard convolution layers with functional modules. Finally, an enhanced module called RCA further refines feature extraction and improves diagnostic performance.

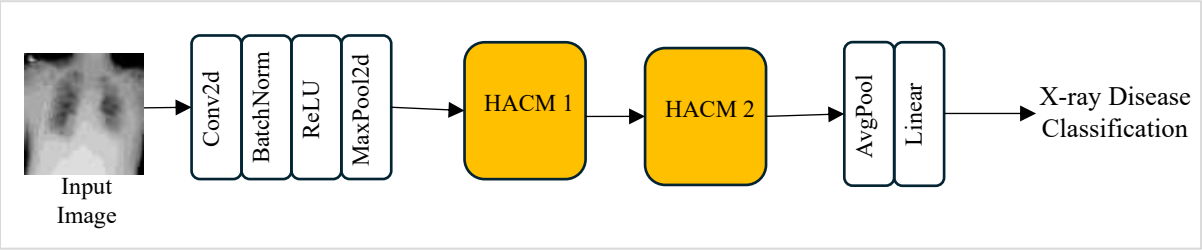


Fig. 2 Structure of the proposed HACNet model.

Geroski et al. [16] explore three weight initialisation schemes (ImageNet, CheXNeXt and DeepCOVID-XR). These schemes are ranked by similarity in image type. ImageNet is the least similar because it was trained on natural images. CheXNeXt is similar, as it was trained on lung diseases other than COVID-19. DeepCOVID-XR is the most similar, as it was trained on COVID-19 X-ray images.

Verma et al. [17] employ a recurrent neural network-based predictor for chest X-ray images to generate artificial visual scan paths. This approach links how humans view images with automated disease classification. They also present a multiclass multilabel classification system that combines scanpaths and image data. This system uses an iterative sequential model with an attention module to enhance classification performance. Their findings show that these simulated human scan paths improve chest X-ray disease recognition.

Hong et al. [18] introduce DTDF-HFF, a deep forest framework based on distance transformation with hybrid feature fusion for accurate chest X-ray image classification. This framework extracts hybrid descriptors from images using handcrafted feature methods and multigrained scanning. It feeds these descriptors into classifiers in each deep forest layer and converts prediction outputs into adaptive distance vectors. These vectors are fused with original features and passed to deeper layers until additional cascades no longer improve accuracy.

Kabi et al. [19] apply the discrete wavelet transform to break chest X-ray images into subbands for multiscale eigendomain gradient boosting (MEGB) to detect pneumonia and TB. For each subband they use singular value decomposition to extract features from singular values and eigenmatrices by selecting the maximum column values. They then concatenate these features and input them into a light gradient boosting model. They validate this approach on a publicly available chest X-ray image dataset.

Wang et al. [20] present SNEL, a stochastic neural ensemble learning setup for robust diagnosis of thoracic diseases from chest X-rays. They construct ensembles of models and employ a noise-robust loss function that enhances both robustness and ensemble diversity to cope with noisy labels. They also use a fast neural ensemble method to aggregate parameters across different model runs and optimisation paths.

2 Proposed Methodology

In this section, we present our novel deep learning architecture for disease classification from X-ray images. The core innovation lies in the development of the Hierarchical Attention-Compression Module (HACM) as shown in Fig. 2, which combines feature aggregation and attentional compression to enhance the model's representational capacity. The following subsections elaborate on the structure of the HACM and its integration into the overall model architecture.

2.1 Hierarchical Attention-Compression Module (HACM)

The Hierarchical Attention-Compression Module (HACM) is designed to process features in a structured and attention-driven manner, enabling the extraction of rich, spatially and channel-wise refined feature representations. It consists of two key components: the *Feature Aggregation Unit* (FAU) and the *Attentional Compression Unit* (ACU).

2.1.1 Feature Aggregation Unit (FAU)

The Feature Aggregation Unit (FAU) is responsible for the dense connectivity of feature maps to promote feature reuse and enhance gradient flow. The FAU combines multiple convolutional layers in a feed forward manner where each layer receives the concatenation of the outputs of all preceding layers as input. This design promotes the generation of highly expressive features by integrating information from multiple scales. Formally, let $x_0, x_1, \dots, x_{l-1}$ represent the outputs of layers 0 to $l-1$, then the input to the l-th layer is:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}])$$

Where $H_l(.)$ represents the composite function (comprising Batch Normalization, ReLU activation, and convolution) applied at the l-th layer, and $[.]$ denotes the concatenation operation. The FAU terminates with a Triplet Attention layer [21], which processes the aggregated features by capturing dependencies across spatial and channel dimensions.

2.1.2 Attentional Compression Unit (ACU)

The Attentional Compression Unit ACU follows the FAU and performs dimensionality reduction while preserving salient features. Unlike traditional downsampling methods the ACU incorporates attention mechanisms for both spatial and channel wise compression. This is achieved through the following stages.

Channel Attention: A squeeze and excitation mechanism evaluates the global importance of each channel by applying global average pooling followed by fully connected layers and a sigmoid activation. The resulting channel weights are used to scale the input feature maps adaptively.

Spatial Attention: A parallel spatial attention mechanism generates an attention map by applying a 2D convolution to the aggregated spatial features. This map highlights the most relevant spatial regions.

Compression: The ACU combines these attention mechanisms with a strided convolution operation ensuring that only the most relevant features are retained during downsampling.

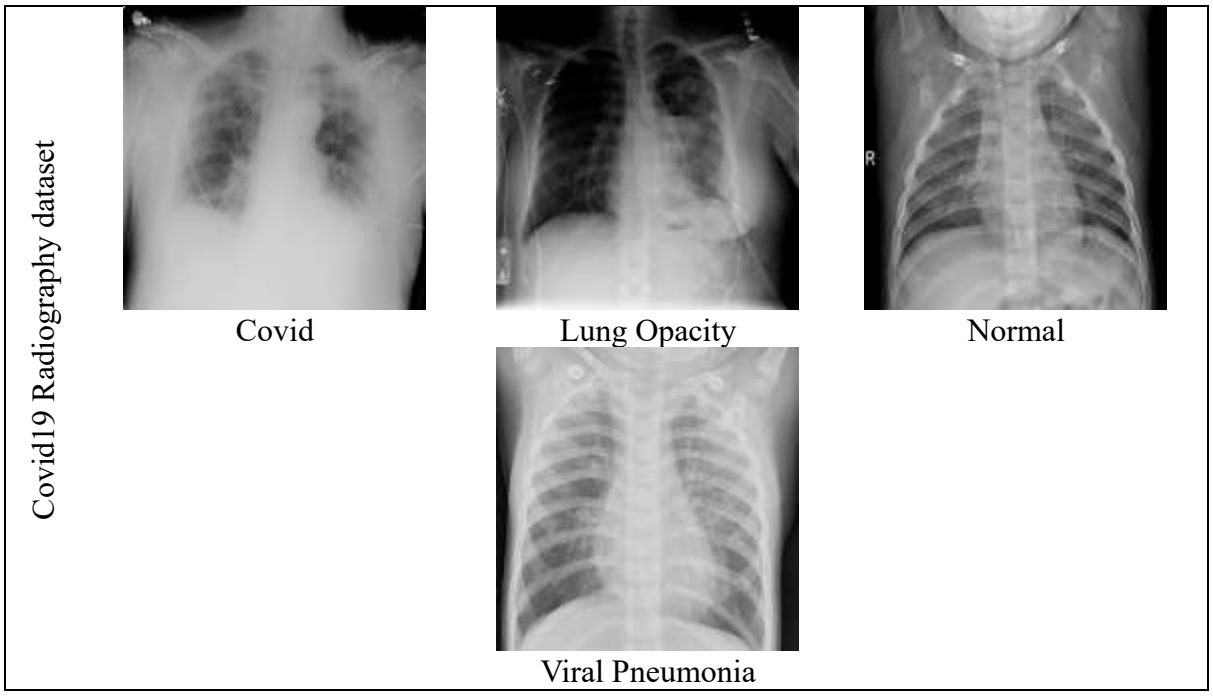
The combined effect of the FAU and ACU ensures hierarchical feature aggregation and compression making HACM an efficient and attention enhanced module.

2.2 Overall Model

The proposed model is a simplified yet powerful design inspired by DenseNet-121 which containing the novel HACM module for enhanced performance. The model is structured as follows:

Initial Layers: The model begins with a sequence of initial layers comprising a convolutional layer with a kernel size of 7×7, stride of 2, and padding, followed by batch normalization, ReLU activation, and a 3×3 max pooling layer.

HACM Blocks: The core of the model consists of **two Hierarchical Attention-Compression Modules (HACMs)**.



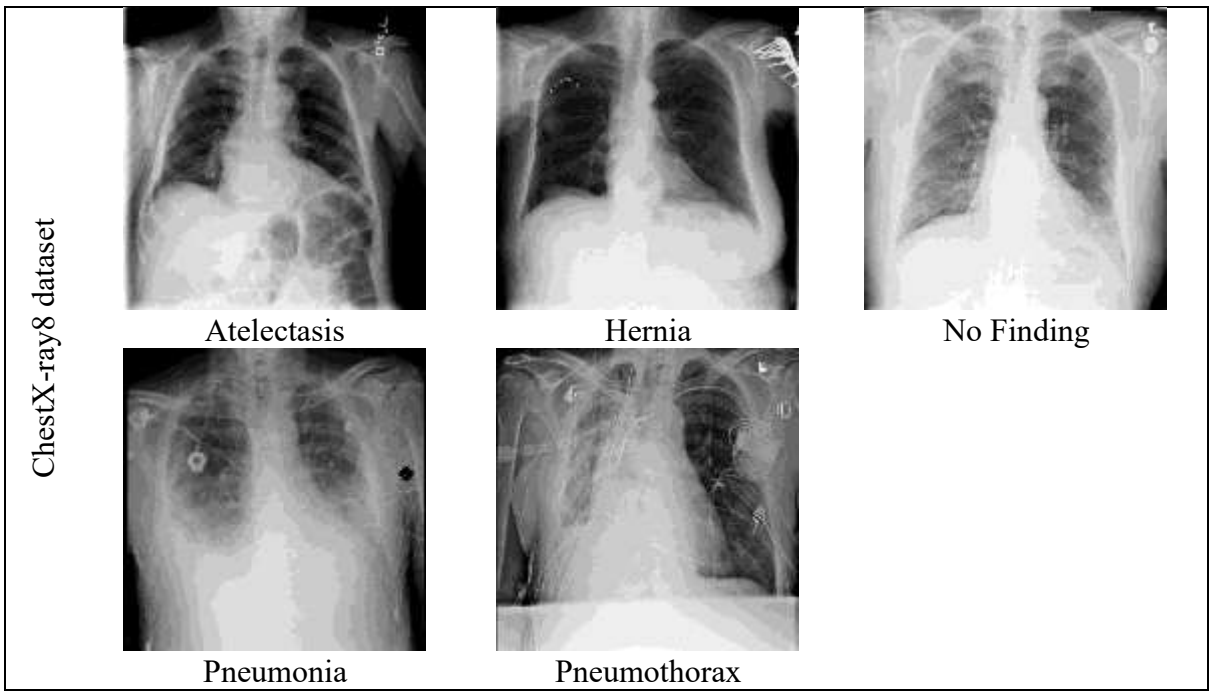


Fig. 3 Dataset images from each class of the Covid19 Radiography and the ChestX-ray8 datasets.

The first HACM processes the output of the initial layers, producing aggregated and compressed features with attention-driven refinement. The second HACM operates on the output of the first HACM and refines the feature representations further.

Final Layers: The HACM blocks first create rich feature maps. A global average pooling layer then squeezes each map down to a single value, giving one long feature vector. This vector moves into a fully connected layer that ends with a softmax, so the network can assign a probability to every disease class and tell which condition an X-ray most likely shows.

During training the model uses categorical cross-entropy to judge its mistakes. Afterward its accuracy, precision, recall, and F1 score show how well it performs. Because HACM captures detailed patterns while still running efficiently, the model offers a practical and powerful tool for medical image analysis.

3 Experimental Setup

3.1 Datasets

Covid19 Radiography dataset [22] contains chest radiograph xray images for machine learning model building and validation to detect COVID19 and associated pulmonary diseases. Images are usually classed as COVID19 positive, normal (healthy), or viral or bacterial pneumonia. The dataset includes patient information, acquisition characteristics, and clinical annotations, and each image is standardized. This form simplifies repeatability and allows thorough comparisons of computational diagnostic methods.

ChestX-Ray8 dataset [23] is a huge dataset of frontal chest radiographs used to build and validate thoracic illness diagnosis algorithms. Over 100,000 photos from over 30,000 patients are labeled for thoracic diseases. The dataset covers diseases like Atelectasis, Hernia, Pneumonia, Pneumothorax and No Finding class for images with no disease. Computer vision

research can perform classification and localization tasks using image-level labels and bounding boxes for a subset of images.

3.2 Preprocessing

All images are resized to 256×256. Augmentations such as random horizontal and vertical flipping are applied to the batch of input images.

3.3 Hyperparameters

The model processes data in batches of 32. It updates its weights with the Adam optimiser. Cross-entropy loss measures how well the classifier is performing. Training starts with a learning rate of 0.0001, and this rate drops by five per percent after every epoch through linear decay. The network trains for 50 epochs on each dataset.

3.4 Hardware

All experiments are conducted on a single NVIDIA T4 GPU having 16GB memory. The system memory is 32GB.

4 Experimental Results & Discussion

This section presents the experimental results and discussions related to the proposed model.

4.1 Performance on Benchmark Datasets

Table 1 shows the performance of the proposed model on the Covid19 radiography and the ChestX-ray8 datasets. Fig. 4 shows these performance scores in the visual form. The proposed model performs well on this classification task.

Table 1 Performance of the proposed model on benchmark datasets.

Datasets	Accuracy	Precision	Recall	F1	AUC	MCC
Covid19 radiography	0.9336	0.9335	0.9336	0.9334	0.8389	0.9056
ChestX-ray8	0.7106	0.7094	0.7106	0.7086	0.6500	0.6389

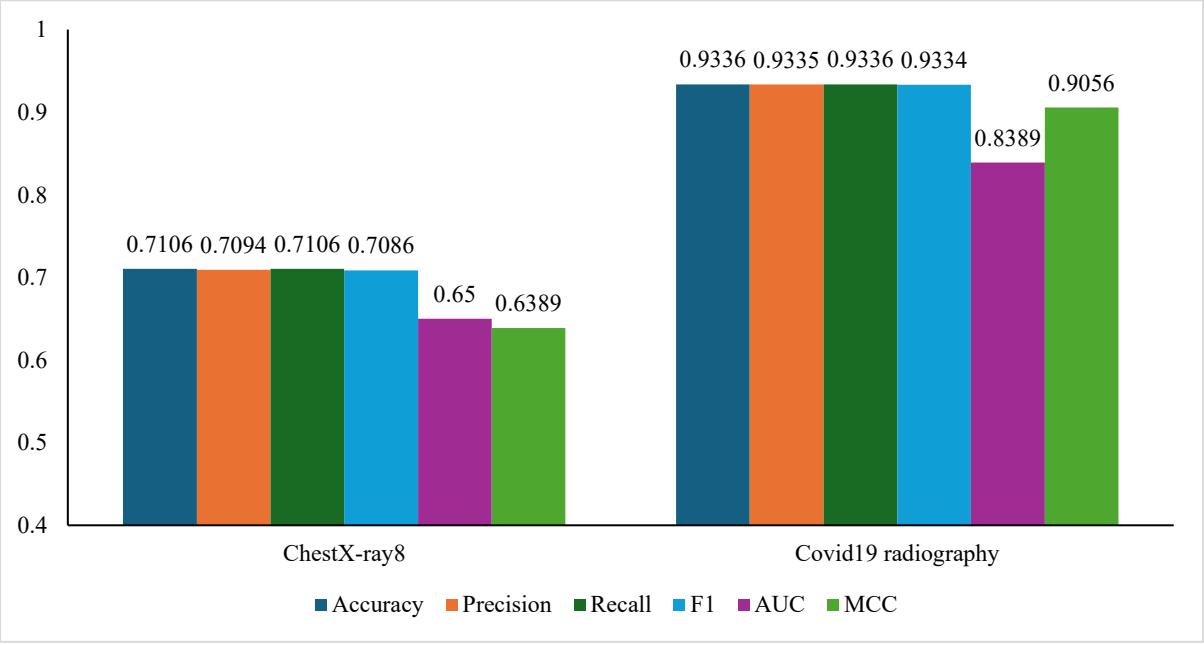


Fig. 4 Performance scores of the proposed model on the ChestX-ray8 and Covid19 radiography dataset.

Fig. 5 and Fig. 6 demonstrate the confusion matrix in identifying various disease from the Covid19 radiography and the ChestX-ray8 datasets. The proposed model is able to identify the correct disease category of most test images.

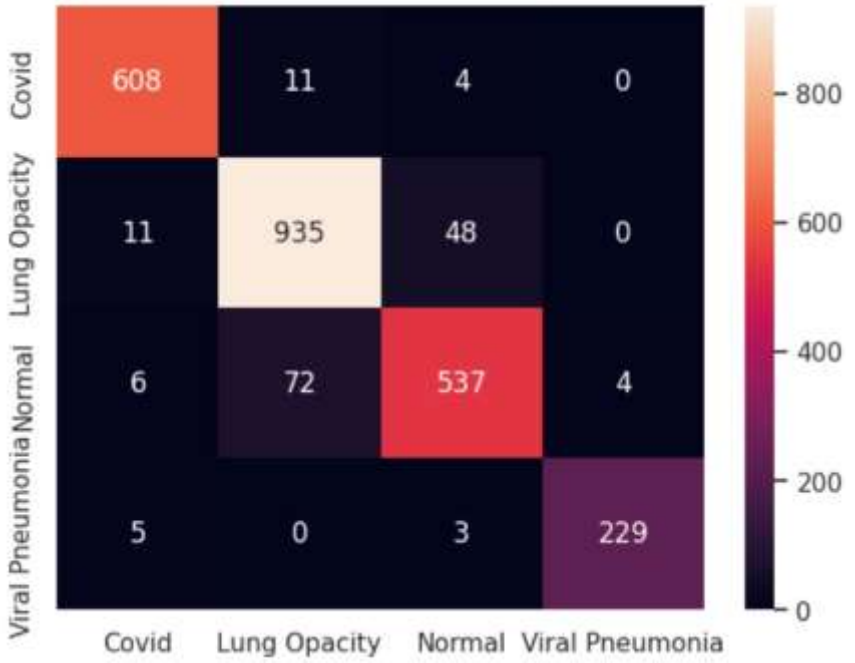


Fig. 5 Confusion matrix of proposed model on the Covid19 radiography dataset.

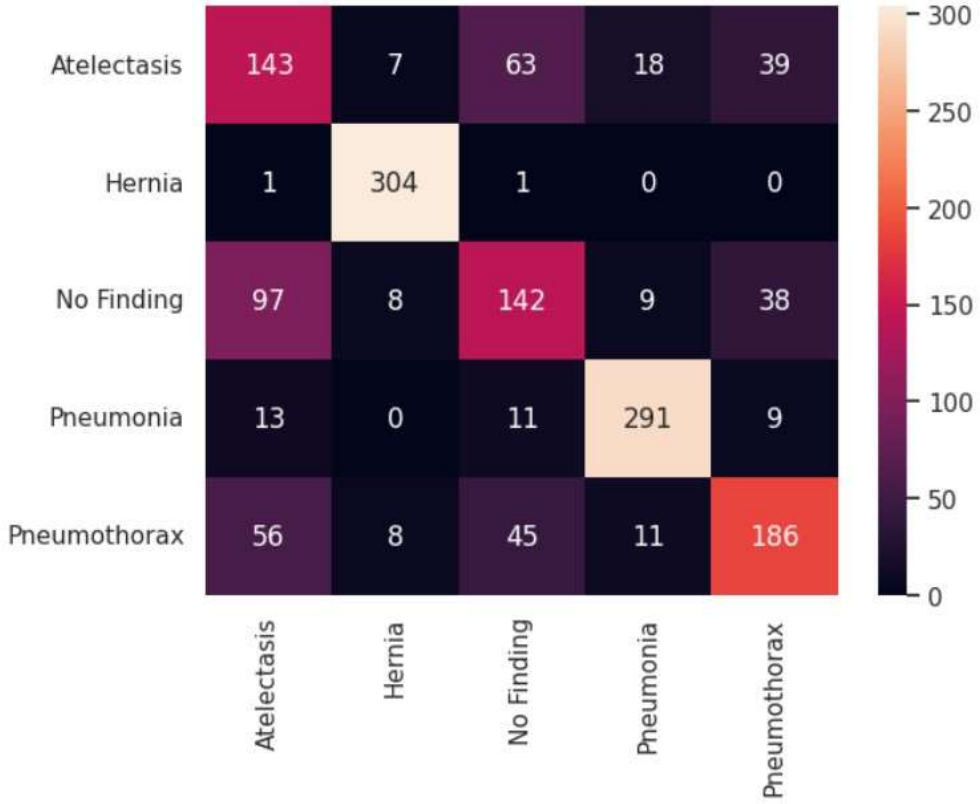
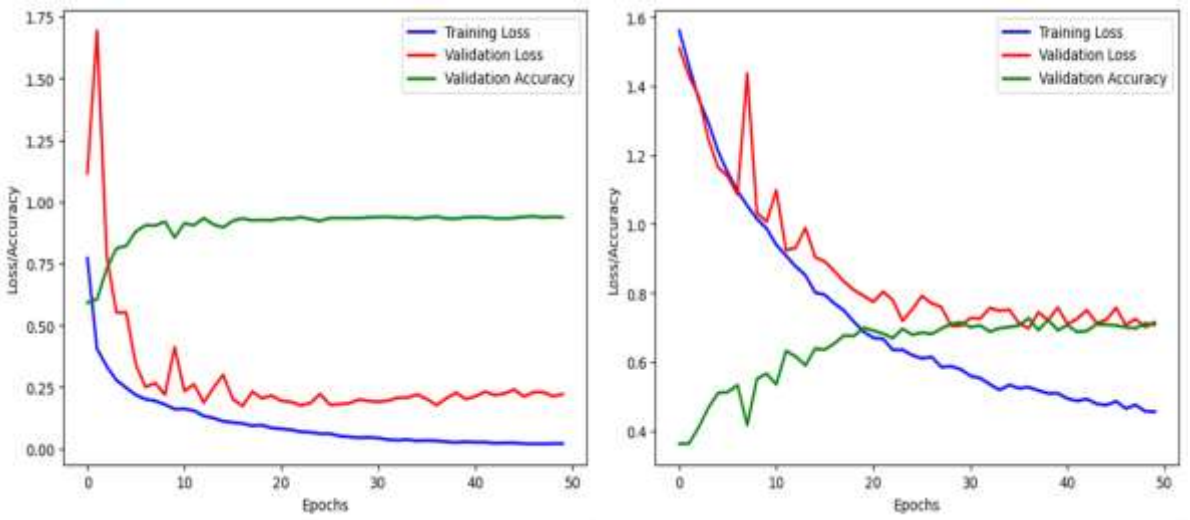


Fig. 6 Confusion matrix of the proposed model on the ChestX-ray8 dataset.

Fig. 7 shows the training loss, validation loss and validation accuracy curves for the proposed model on the Covid19 radiography and the ChestX-ray8 dataset. The model was trained for 25 epochs for each dataset.



a) Covid19 radiography b) ChestX-ray8

Fig. 7 Training, validation loss and validation accuracy curves for the proposed model on a) Covid19 radiography dataset and b) ChestX-ray8 dataset.

4.2 Complexity Analysis

This part measures how demanding the new model is and compares it with leading computer-vision networks. It reviews four points: how many parameters the model has, how many MAC operations it needs, and how long it takes to run on a CPU or a GPU. The new model is very light because it uses only 1.43 million parameters.

Table 2 Complexity analysis of the proposed model.

Model	Parameter (millions)	MACs ($\times 10^9$)	CPU time (s)	GPU time (s)
ConvNeXt	88.59	5.02	4.95	1.18
EfficientNet	66.34	1.75	7.27	1.43
EfficientNet v2	118.51	4.05	6.61	0.96
Swin Transformer	87.93	3.38	6.69	0.93
ResNet152	60.19	3.79	4.41	0.32
ResNeXt	88.79	5.40	4.75	0.87
HACNet (proposed)	1.43	2.08	1.46	0.28

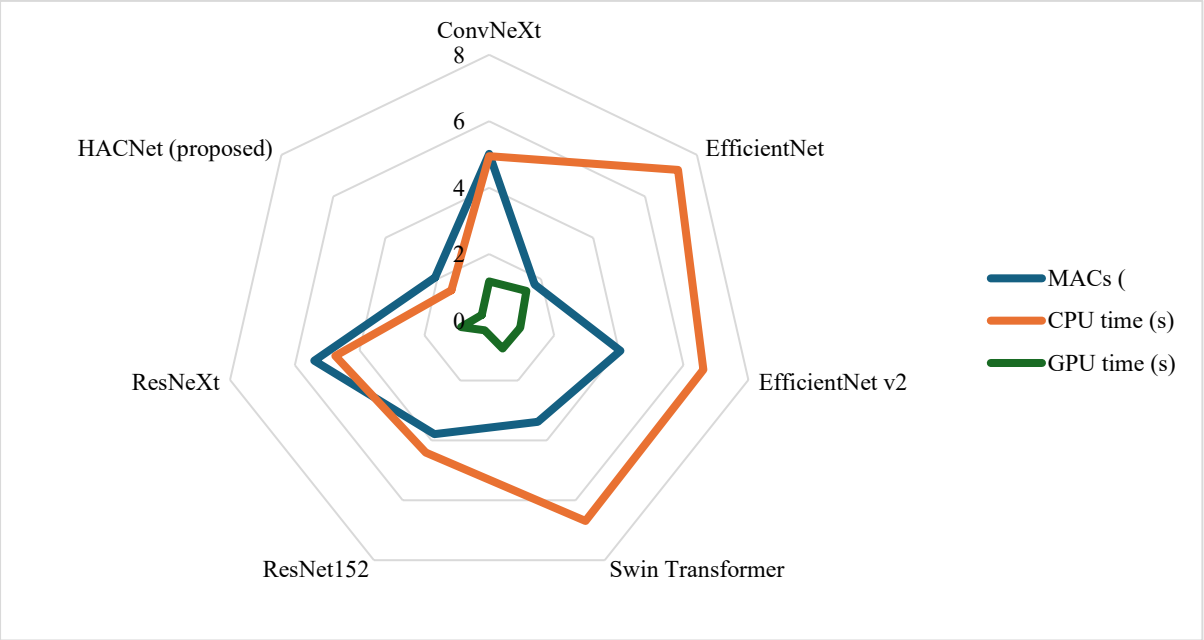


Fig. 8 Computational comparison of proposed model against existing state-of-the-arts in computer vision. HACNet has very low computational requirements compared to popular computer vision models.

Fig. 8 clearly proves that the proposed model is computationally lightweight and effective. It has a low MAC count, and its CPU and GPU inference time is considerably low compared to popular computer vision models.

4.3 Qualitative Analysis

This section presents the qualitative analysis of the proposed model. Specifically, it presents the t-SNE visualization of the features extracted by the proposed model as shown in Fig. 9 and Fig. 10. The data points from each class are clearly differentiated based on the extracted features indicating strong classification capabilities of the proposed model.

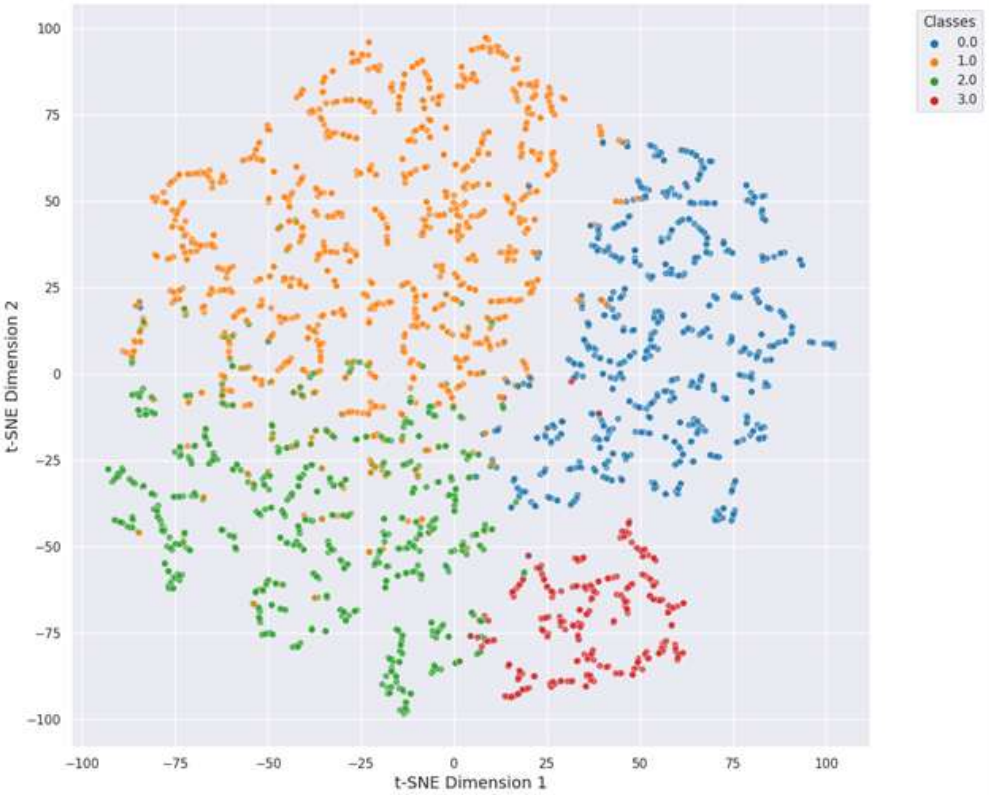


Fig. 9 t-SNE plot of features from the proposed model on the Covid19 radiography dataset.

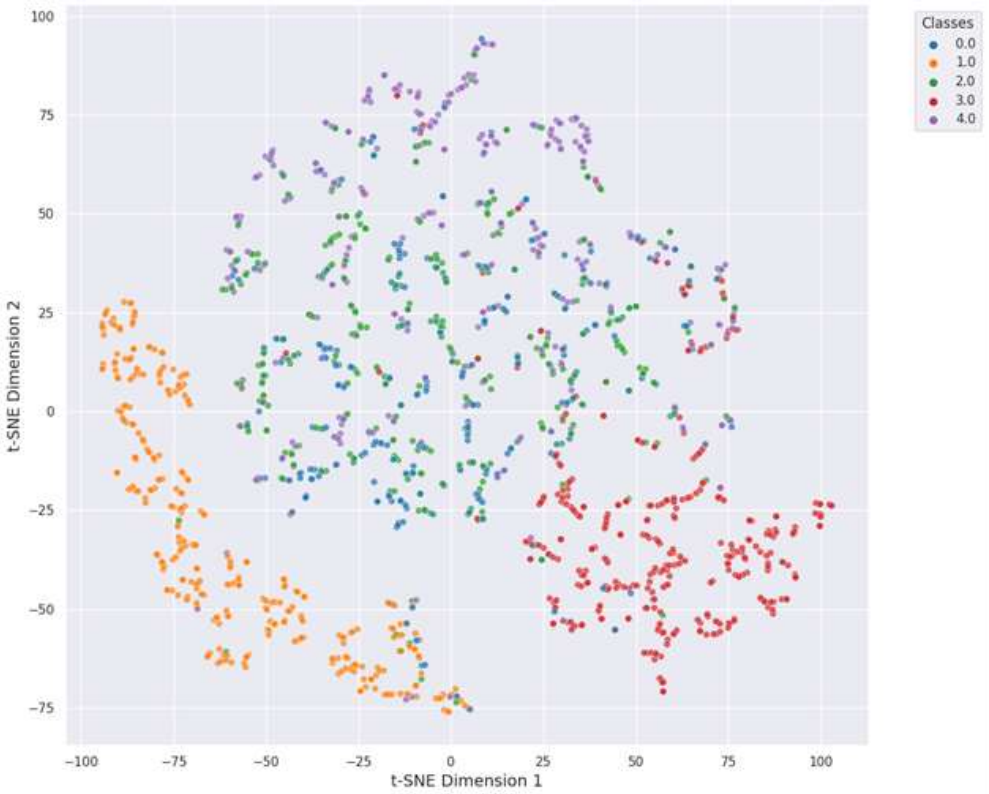


Fig. 10 t-SNE plot of features from the proposed model on the ChestX-ray8 dataset.

5 Conclusion, Limitations & Future Work

5.1 Conclusion

We presented HACNet, a deep learning framework for disease detection from x-ray images that combines our Hierarchical Attention-Compression Module (HACM) into a simplified network design. The Feature Aggregation Unit (FAU) and Attentional Compression Unit (ACU) synergistically collect rich multi-scale, spatial, and channel-wise characteristics in HACNet, creating an expressive and computationally efficient model. HACNet has shown promise in improving automated disease diagnosis by adding sophisticated attention mechanisms.

5.2 Limitations

HACNet's promising performance needs more examination due to few possible constraints. The model's generalization to new clinical circumstances is limited by the quality and variety of the training data. Image capture conditions or a shortage of samples could impact this. The HACM balances feature aggregation with attention-driven compression, but its architectural complexity may increase computational cost during training, making real-time implementation on resource-constrained devices difficult.

5.3 Future Work

Future research should build on HACNet's strengths to improve the model's generalization across diverse datasets [24] by using unsupervised and semi-supervised learning techniques to reduce the need for extensive annotated data. Implementing model pruning, quantization, or knowledge distillation can be used to increase computing performance and allow for real-time deployment. HACNet might also be expanded to include multi-modal data inputs including dermoscopic pictures and patient information for a more complete and context-aware diagnosis tool. These directions might improve clinical decision-making and automated disease diagnosis [25].

Data Availability, Funding and Declaration Statements

Data Availability

The datasets produced and/or analysed during the current study are available via an online web repository, which may be accessed via the weblinks provided below:

Covid19 Radiography database - https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database

NIH Chest-xrays - https://www.kaggle.com/datasets/nih-chest-xrays/data

Funding

No funds were received for this manuscript.

Declarations

There are no known personal ties or conflicting financial interests between the authors that may affect the work presented in this study.

6 References

[1] M. Kaya and Y. Çetin-Kaya, "A novel ensemble learning framework based on a genetic algorithm for the classification of pneumonia," *Eng. Appl. Artif. Intell.*, vol. 133, p. 108494, Jul. 2024, doi: 10.1016/j.engappai.2024.108494.

[2] I. Adjei-Mensah *et al.*., "Cov-Fed: Federated learning-based framework for COVID-19 diagnosis using chest X-ray scans," *Eng. Appl. Artif. Intell.*, vol. 128, p. 107448, Feb. 2024, doi: 10.1016/j.engappai.2023.107448.

[3] P. K. Das, S. Sreevatsav, and A. Abraham, "An efficient deep learning network with orthogonal softmax layer for automatic detection of tuberculosis," *Eng. Appl. Artif. Intell.*, vol. 133, p. 108116, Jul. 2024, doi: 10.1016/j.engappai.2024.108116.

[4] M. Irfan, K. M. Malik, and K. Muhammad, "Federated fusion learning with attention mechanism for multi-client medical image analysis," *Inf. Fusion*, vol. 108, p. 102364, Aug. 2024, doi: 10.1016/j.inffus.2024.102364.

[5] J. Rocha, S. C. Pereira, J. Pedrosa, A. Campilho, and A. M. Mendonça, "STERN: Attention-driven Spatial Transformer Network for abnormality detection in chest X-ray images," *Artif. Intell. Med.*, vol. 147, p. 102737, Jan. 2024, doi: 10.1016/j.artmed.2023.102737.

[6] Z. Ren, S. Liu, L. Wang, and Z. Guo, "Conv-SdMLPMixer: A hybrid medical image classification network based on multi-branch CNN and multi-scale multi-dimensional MLP," *Inf. Fusion*, vol. 118, p. 102937, Jun. 2025, doi: 10.1016/j.inffus.2025.102937.

[7] A. Yadav and D. K. Vishwakarma, "MRT-Net: Auto-adaptive weighting of manipulation residuals and texture clues for face manipulation detection," *Expert Syst. Appl.*, vol. 232, p. 120898, Dec. 2023, doi: 10.1016/j.eswa.2023.120898.

[8] A. B. Subba and A. K. Sunaniya, "Computationally optimized brain tumor classification using attention based GoogLeNet-style CNN," *Expert Syst. Appl.*, vol. 260, p. 125443, Jan. 2025, doi: 10.1016/j.eswa.2024.125443.

[9] L. Mei *et al.*., "Clinical knowledge-guided hybrid classification network for automatic periodontal disease diagnosis in X-ray image," *Med. Image Anal.*, vol. 99, p. 103376, Jan. 2025, doi: 10.1016/j.media.2024.103376.

[10] Ş. Öztürk, M. Y. Turalı, and T. Çukur, "HydraViT: Adaptive multi-branch transformer for multi-label disease classification from Chest X-ray images," *Biomed. Signal Process. Control*, vol. 100, p. 106959, Feb. 2025, doi: 10.1016/j.bspc.2024.106959.

[11] J. P. G, P. S, V. V. Mayil, and S. Saini, "Optimized double transformer residual super-resolution network-based X-ray images for classification of pneumonia identification," *Knowl.-Based Syst.*, vol. 311, p. 113037, Feb. 2025, doi: 10.1016/j.knosys.2025.113037.

[12] X. Zhao and X. Wang, "Multi-label chest X-ray image classification based on long-range dependencies capture and label relationships learning," *Biomed. Signal Process. Control*, vol. 100, p. 107018, Feb. 2025, doi: 10.1016/j.bspc.2024.107018.

[13] M. Patankar, V. Chaurasia, and M. Shandilya, "A novel spiking neural network method for classification of tuberculosis using X-ray images," *Comput. Electr. Eng.*, vol. 122, p. 110003, Mar. 2025, doi: 10.1016/j.compeleceng.2024.110003.

[14] J. Rabbah, M. Ridouani, and L. Hassouni, "Improving pneumonia diagnosis with high-accuracy CNN-Based chest X-ray image classification and integrated gradient," *Biomed. Signal Process. Control*, vol. 101, p. 107239, Mar. 2025, doi: 10.1016/j.bspc.2024.107239.

[15] S. Yi, S. Qin, F. She, and D. Shao, "BSD: A multi-task framework for pulmonary disease classification using deep learning," *Expert Syst. Appl.*, vol. 259, p. 125355, Jan. 2025, doi: 10.1016/j.eswa.2024.125355.

[16] T. Geroski *et al.*., "Enhancing COVID-19 disease severity classification through advanced transfer learning techniques and optimal weight initialization schemes," *Biomed. Signal Process. Control*, vol. 100, p. 107103, Feb. 2025, doi: 10.1016/j.bspc.2024.107103.

[17] A. Verma, A. Kar, K. Ghosh, S. K. Dhara, D. Sen, and P. Kumar Biswas, "Artificially Generated Visual Scanpath Improves Multilabel Thoracic Disease Classification in Chest X-Ray Images," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–11, 2024, doi: 10.1109/TIM.2024.3428591.

[18] Q. Hong *et al.*., "A Distance Transformation Deep Forest Framework With Hybrid-Feature Fusion for CXR Image Classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 10, pp. 14633–14644, Oct. 2024, doi: 10.1109/TNNLS.2023.3280646.

[19] S. K. Kabi, R. K. Tripathy, D. Patra, and G. Panda, "A Novel Approach for the Detection of Tuberculosis and Pneumonia Using Chest X-Ray Images for Smart Healthcare Applications," *IEEE Sens. Lett.*, vol. 7, no. 12, pp. 1–4, Dec. 2023, doi: 10.1109/LSENS.2023.3327580.

[20] "Robust Stochastic Neural Ensemble Learning With Noisy Labels for Thoracic Disease Classification | IEEE Journals & Magazine | IEEE Xplore." Accessed: Feb. 07, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/10413624

[21] D. Misra, T. Nalamada, A. U. Arasanipalai, and Q. Hou, "Rotate to Attend: Convolutional Triplet Attention Module," in *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Jan. 2021, pp. 3138–3147. doi: 10.1109/WACV48630.2021.00318.

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

[23] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 3462–3471. doi: 10.1109/CVPR.2017.369.

[24] L. Zou *et al.*., "MCG-Net: Medical Chief Complaint-guided Multi-modal Masked Content Pre-training for chest image classification," *Expert Syst. Appl.*, vol. 271, p. 126660, May 2025, doi: 10.1016/j.eswa.2025.126660.

[25] T. Gulsoy and E. Baykal Kablan, "FocalNeXt: A ConvNeXt augmented FocalNet architecture for lung cancer classification from CT-scan images," *Expert Syst. Appl.*, vol. 261, p. 125553, Feb. 2025, doi: 10.1016/j.eswa.2024.125553.