Open Access and Peer Review Journal ISSN 2394-2231 https://ijctjournal.org/

# CAXNet: Coordinate Attention-Enhanced ResNeXt with Multi-Level Feature Fusion for Multiclass Chest X-Ray Classification

Akshay Mool, Ashutosh Pandey, Rahul Kumar, Ankit Yadav
*Assistant Professor, Delhi Technological University*

----------**********----------

**Abstract**—Rapid and reliable COVID-19 detection from chest X-rays can aid clinical screening and radiologists' decisions. Due to overlapping radiographic features in COVID-19, lung opacity, viral pneumonia, and normal chest X-rays, multiclass categorization is difficult. This study offers CAXNet, a Coordinate Attention-enhanced ResNeXt COVID-19 chest X-ray classification model. This architecture enhances ResNeXt50-32x4d by preserving early and intermediate residual stages and incorporating Coordinate Attention modules for improved spatial feature representations. Before categorization, attention-enhanced low- and intermediate-level characteristics are fused using multi-level feature fusion. The model was tested using the COVID-19 Radiography Database for four-class classification: disease, lung opacity, normal, and viral pneumonia. The experimental findings demonstrate that CAXNet exhibited 92.00\% accuracy, 92.15\% precision, 92.00\% recall, 92.00\% F1-score, 0.8852 Matthews Correlation Coefficient, and 0.9329 ROC-AUC on the test set. The confusion matrix and t-SNE-based feature visualization show that the proposed model learns discriminative representations, notably for COVID-19 and Viral Pneumonia. The results show that coordinate-aware attention and multi-level feature fusion enhance ResNeXt-based chest X-ray image classification models.

**Index Terms**—COVID-19 classification; chest X-ray; deep learning; ResNeXt; Coordinate Attention; feature fusion; medical image classification

----------**********----------

## I. INTRODUCTION

COVID-19 has burdened global healthcare systems and raised the demand for fast, reliable, and affordable diagnostic help. Lung involvement requires radiological proof, hence medical imaging is crucial in respiratory infection diagnosis. Many healthcare facilities utilize chest X-ray imaging because it is cheaper, faster, and more accessible than computed tomography. Thus, automated chest X-ray analysis is a key COVID-19 screening and pulmonary illness categorization study area.

Deep learning approaches, notably convolutional neural networks, appear promising for medical picture categorization. Several studies have improved COVID-19 chest X-ray categorization using pretrained CNN architectures, transfer learning, tailored CNN models, feature fusion, and attention techniques. Transfer learning-based approaches are effective when medical imaging data is scarce, whereas specialized deep structures can capture disease-specific radiographic patterns. Different lung illnesses can create visually similar patterns, making multiclass chest X-ray categorization difficult. COVID-19, Lung Opacity, and Viral Pneumonia may share radiographic opacity areas and textural differences, making class differentiation difficult.

For classification, most approaches employ entire pretrained CNN backbones. Such models have excellent feature extraction, but they may also include deeper layers tuned for large-scale natural picture identification rather than medical radiography analysis. Low-level and intermediate-level feature maps may show disease-specific texture, edge, opacity, and localized structural abnormalities in chest X-ray categorization. Thus, a good model should conserve useful radiographic characteristics while learning abstract categorization representations.

Deep learning models use attention techniques to highlight essential regions and suppress irrelevant information to enhance feature representation. Disease symptoms are generally spatially confined, making medical imaging attention valuable. Global pooling of channel descriptors may cause traditional attention techniques to lose positional information. Coordinate Attention encodes height and breadth spatial information to preserve positional cues and describe channel-wise interdependence. For chest X-ray categorization, the geographical distribution of aberrant spots can help classify illness types.

CAXNet, a Coordinate Attention-enhanced ResNeXt architecture for COVID-19 chest X-ray classification, addresses these concerns. This model alters the ResNeXt50-32x4d backbone by keeping only the stem block and the first two residual stages. After maintained residual stages, Coordinate Attention modules enhance derived feature maps. A multi-level feature fusion approach fuses attention-enhanced features from distinct stages. The model may integrate fine-grained local radiography patterns with deeper semantic representations. Finally, global average pooling and a fully linked classifier classify four classes.

The CAXNet model classifies chest X-ray pictures into COVID-19, Lung Opacity, Normal, and Viral Pneumonia on the COVID-19 Radiography Database. Experimental findings show that the suggested architecture performs well in classification and learns meaningful class-specific representations. The confusion matrix and t-SNE visualization help the feature extraction and fusion technique work.

Summary of this work's main contributions:

- CAXNet, a modified ResNeXt50-32x4d architecture, is presented for multiclass COVID-19 chest X-ray classification.

- Adapting the ResNeXt50-32×4d backbone to maintain just early and intermediate residual stages results in a more compact design while keeping radiographic feature representations.
- Integrate Coordinate Attention modules to improve disease-relevant feature maps and increase spatial information.
- A multi-level feature fusion technique combines attention-refined low- and intermediate-level characteristics for categorization.
- The model performs well on a four-class chest X-ray classification job, with 92.00% accuracy, 92.05% weighted F1-score, 0.8852 MCC, and 0.9329 weighted ROC-AUC.

## II. RELATED WORK

Automatic COVID-19 identification and categorization using deep learning-based chest X-ray analysis is an area which is widely researched. Later research classified normal, pneumonia, lung opacity, and other thoracic diseases into multiclass groups, while early investigations concentrated on binary categorization. Chest X-ray imaging is cheap, rapid, and widely available, making it helpful for building computer-aided diagnostic systems.

Ohata et al. [1] suggested a transfer learning-based chest X-ray COVID-19 detection framework. Pretrained convolutional neural networks extracted features and were integrated with standard machine learning classifiers. Transfer learning worked even with insufficient COVID-19 picture data. Khan et al. [2] introduced CoroNet, a CNN-based architecture employing the Xception model to diagnose COVID-19 from chest X-rays. Their model showed that pretrained deep CNNs can classify radiographic diseases categorized into COVID-19, pneumonia, and normal classes.

Several research suggest customized CNN architectures for COVID-19 and pneumonia classification. The multi-dilation CNN model - CovXNet, by Mahmud et al. [3] detects COVID-19 and pneumonia from chest X-rays automatically. The model captured characteristics from distinct receptive fields using multi-dilation convolutions. Abbas et al. [4] introduced DeTraC, a deep convolutional neural network for COVID-19 classification using decomposition, transfer, and composition. Class decomposition was used to resolve data distribution inconsistencies and improve chest X-ray feature separability. Jain et al. [5] examined deep CNN models for COVID-19 detection and analysis using chest X-ray images and stressed the need of preprocessing, augmentation, and model comparison in diagnostic systems.

Image preprocessing and augmentation have also been studied in COVID-19 chest X-ray categorization. Rahman et al. [6] examined image enhancing methods for COVID-19 detection in chest X-rays. They found that enhancement approaches improve disease-relevant pattern visibility and classification performance. This is essential since chest X-ray datasets generally come from diverse sources with different contrast, brightness, and acquisition settings.

Deep learning models' discrimination has been improved via feature fusion. Kong and Cheng [7] presented DenseNet and VGG16 feature fusion for COVID-19 X-ray image categorization. Their method incorporated complementing deep features from two CNN architectures. The explainable deep learning-based model by Bhandari et al. [8] classified chest X-ray pictures into COVID-19, pneumonia, TB, and normal classes. Their work focused on classification performance and explainability, emphasizing interpretable deep learning in medical picture analysis.

Other research have examined whether fine-tuning pretrained CNNs or designing new models is essential. Pham et. al. [9] used deep learning to classify COVID-19 chest X-rays and compared new models to fine-tuned pretrained networks. When datasets are limited, fine-tuned pretrained models can compete for COVID-19 classification, according to the study. Oltu et al. [10] recently suggested an attention-based deep learning solution for automated COVID-19, lung opacity, and viral pneumonia chest X-ray categorization. Advanced feature extraction and attention techniques make their work useful to multiclass COVID-19 radiographic classification.

The literature suggests that transfer learning, custom CNN design, preprocessing, feature fusion, and attention mechanisms improve COVID-19 chest X-ray categorization. Many previous approaches employ entire pretrained backbones directly, use multiple-network fusion, or apply attention without explicitly integrating coordinate-aware refinement and multi-level feature fusion. The CAXNet model changes the ResNeXt50-32×4d backbone by maintaining just early and intermediate residual stages, integrating Coordinate Attention modules, and fusing attention-refined multi-level feature maps before categorization. This design preserves local radiographic information and captures intermediate semantic representations for multiclass COVID-19 chest X-ray classification.

## III. PROPOSED METHODOLOGY

### *A. Overview of the Proposed CAXNet Architecture*

CAXNet, a Coordinate Attention-enhanced ResNeXt architecture, is proposed for multiclass COVID-19 chest X-ray classification. The suggested model classifies chest X-rays into COVID-19, Lung Opacity, Normal, and Viral Pneumonia. To generate the model, the ResNeXt50-32x4d architecture was modified and Coordinate Attention modules were integrated with a multi-level feature fusion technique.

The input to the proposed model is a grayscale chest X-ray image resized to 256×256. Therefore, each input image can be represented as:

$$X \in \mathbb{R}^{1\times256\times256}$$

As the ResNeXt50-32×4d model was built for three-channel RGB pictures, its first convolutional layer was changed to take a single-channel grayscale input. CAXNet only uses the first stem block and the first two residual stages, instead of the full ResNeXt50-32×4d design. The initial categorization layer and deeper residual stages are deleted.

This architectural change reduces model size while keeping low- and intermediate-level radiography properties.

These traits help classify chest X-rays because disease-related patterns generally present as local texture changes, opacity patches, and lung structural anomalies.

A modified ResNeXt-based feature extractor, Coordinate Attention modules for feature refinement, and a multi-level feature fusion-based classification head make up CAXNet.

### *B. Coordinate Attention-Enhanced ResNeXt Feature Extraction*

CAXNet's feature extraction component uses the ResNeXt50-32x4d backbone. Because aggregated residual modifications increase feature representation, ResNeXt is useful for picture classification. The suggested approach retains just the early and middle ResNeXt phases to extract important radiography characteristics while minimizing architectural depth.

Let S(·) denote the modified stem block of ResNeXt, and let $R_1$ (·) and $R_2$ (·) denote the first and second residual stages, respectively. The feature maps extracted by these stages can be represented as:

$$F_1 = R_1(S(X)), F_2 = R_2(F_1)$$

Here, $F_1$ represents the feature map obtained from the first residual stage, while $F_2$ represents the feature map obtained from the second residual stage. A 1 × 256 × 256 input picture yields a 256 × 64 × 64 feature map, whereas a 512 × 32 × 32 feature map is produced by the second residual stage.

Coordinate Attention modules after both residual phases increase feature map discrimination. Coordinate Attention effectively captures channel-wise dependencies while retaining height and breadth positional information. This is important for chest X-ray categorization, since illness symptoms may be concentrated in certain lung areas.

The attention-refined feature maps are obtained as:

$$F_1 = CA_1(F_1), F_2 = CA_2(F_2)$$

where $CA_1(\cdot)$ and $CA_2(\cdot)$ denote the Coordinate Attention modules applied after the first and second ResNeXt residual stages, respectively. The initial attention-refined feature map $F_1$ includes details like edges, textures, and local opacity patterns. $F_2$ is a feature map that captures more abstract and detailed disease-related representations. In this way, CAXNet integrates Coordinate Attention into the updated ResNeXt backbone to enhance features.

### *C. Multi-Level Feature Fusion and Classification Head*

The feature maps from the two residual stages are combined after attention refinement to provide a richer representation. $F_1$ and $F_2$ have different spatial dimensions, therefore the first-stage attention-refined feature map is downsampled before fusing. The fusion operation is defined as:

$$F_{\text{fused}} = Concat(D(F_1), F_2)$$

where *D(.)* denotes downsampling and *Concat(.)* denotes channel-wise concatenation. After downsampling, $F_1$ and $F_2$ become spatially consistent. The fused feature map has 768 channels by concatenating 256 channels from the first residual stage and 512 from the second.

This multi-level fusion technique lets the model incorporate complementing information from two representation levels. First-stage features maintain local and fine-grained radiography patterns, whereas second-stage features give abstract and semantically relevant information. The classifier uses their fusion to differentiate visually similar classes like COVID-19, Lung Opacity, and Viral Pneumonia.

Global average pooling creates a compact feature vector from the fused feature map. This vector is delivered to a fully connected layer, followed by softmax for classification:

$$\hat{y} = Softmax(W \cdot GAP(F_{\text{fused}}) + b)$$

where *GAP(·)* denotes global average pooling, *W* and *b* are the trainable parameters of the fully connected layer, and $\hat{y}$ represents the predicted class probability distribution.

The final predicted class is obtained as:

$$\hat{c} = \arg\max_k \hat{y}_k$$

where $\hat{c}$ is the predicted class label and $\hat{y}_k$ is the predicted probability of the $k^{th}$class.

Therefore, the proposed CAXNet architecture includes a compact ResNeXt-based feature extractor, Coordinate Attention-based feature refining, and multi-level feature fusion. This allows the model to classify COVID-19 chest X-rays using local radiographic information and disease-specific representations.

## IV. EXPERIMENTAL SETUP

### *A. Dataset Description*

The COVID-19 Radiography Database was used to test the propoed CAXNet model on chest X-ray pictures from four categories: COVID-19, Lung Opacity, Normal, and Viral Pneumonia. The multiclass classification job targeted these four groups.

Each chest X-ray image was assigned one of the following labels:

**TABLE I**
**DATASET CLASS LABELS**

| Label | Class |
|---|---|
| 0 | COVID-19 |
| 1 | Lung Opacity |
| 2 | Normal |
| 3 | Viral Pneumonia |

A total of 6,500 images were in the experiment. The data was split into training, validation, and testing. Training set optimized model parameters, validation set monitored learning behavior during training, and test set evaluated final performance.

**TABLE II**
**DATA SPLIT AND IMAGE COUNTS**

| Subset | Number of images |
|---|---|
| Training set | 5,000 |
| Validation set | 500 |
| Test set | 1,000 |

| | |
|---|---|
| Total | 6,500 |

### *B. Data Preprocessing and Augmentation*

Each chest X-ray image was converted to grayscale and shrunk to a fixed 256×256 spatial resolution. The first convolutional layer of ResNeXt50-32×4d was updated to take single-channel grayscale input, while the original architecture required three-channel RGB pictures.

The following preprocessing and augmentation operations were applied:

**TABLE III**
**IMAGE PREPROCESSING AND AUGMENTATION OPERATIONS**

| Operation | Description |
|---|---|
| Grayscale conversion | Each image was converted into a single-channel image. |
| Resizing | Images were resized to 224 × 224 pixels. |
| Random horizontal flipping | Used for data augmentation. |
| Random vertical flipping | Used for data augmentation. |
| Tensor conversion | Images were converted into PyTorch tensors. |

Data augmentation increased training sample variability and reduced overfitting. Augmentation can increase model resilience in chest X-ray datasets with variability in acquisition circumstances, image orientation, and patient posture.

### *C. Implementation Details*

PyTorch deep learning implemented the proposed CAXNet model. The pretrained ResNeXt50-32×4d model was used to initialize the backbone network. For grayscale chest X-ray input, the first convolutional layer was substituted with a single-channel one. The last classification layer and deeper residual stages of ResNeXt50-32×4d were eliminated. Integrating Coordinate Attention modules and the multi-level feature fusion block into the updated backbone.

The Adam optimizer with 0.001 learning rate was used to trained the model. The categorical cross-entropy loss function was utilized for the four-class classification issue. StepLR learning rate schedulers with decay factors of 0.9 were implemented after each epoch. The model was trained for 20 epochs with 64 batches.

The main hyperparameters used in the experiment are summarized below:

**TABLE IV**
**EXPERIMENTAL HYPERPARAMETERS**

| Parameter | Value |
|---|---|
| Input image size | 224 × 224 |
| Input channels | 1 |
| Number of classes | 4 |
| Batch size | 64 |
| Number of epochs | 20 |
| Optimizer | Adam |
| Initial learning rate | 0.001 |
| Learning rate scheduler | StepLR |
| Scheduler gamma | 0.9 |
| Loss function | Cross-entropy loss |
| Backbone | ResNeXt50-32×4d |
| Attention mechanism | Coordinate Attention |

### *D. Evaluation Metrics*

Several standard classification metrics were used to analyze CAXNet. The percentage of correctly categorized photos was measured by accuracy. Weighted averaging was used to calculate precision, recall, and F1-score for test set class imbalance. Matthews Correlation Coefficient was also presented since it balances categorization quality by considering true positives, true negatives, false positives, and false negatives. In addition, weighted multiclass ROC-AUC assessed the model's class-separation performance.

The following metrics were used for final evaluation:

**TABLE V**
**EVALUATION METRICS AND PURPOSES**

| Metric | Purpose |
|---|---|
| Accuracy | Measures overall classification correctness. |
| Weighted precision | Measures class-balanced predictive reliability. |
| Weighted recall | Measures class-balanced sensitivity. |
| Weighted F1-score | Harmonic mean of weighted precision and recall. |
| MCC | Measures overall prediction-ground truth agreement. |
| ROC-AUC | Measures multiclass separability. |
| Confusion matrix | Shows class-wise correct and incorrect predictions. |
| t-SNE visualization | Provides qualitative feature-space analysis. |

A confusion matrix identified the main causes of misclassification among the four classifications. Also, t-SNE visualization was created from the 768-dimensional penultimate feature vectors after global average pooling. This visualization assessed whether learnt feature representations formed class-specific clusters qualitatively.

## V. RESULTS AND DISCUSSIONS

Experimental findings from the proposed CAXNet model on COVID-19 chest X-ray categorization are presented here. Accuracy, precision, recall, F1-score, Matthews Correlation Coefficient, ROC-AUC, confusion matrix, and t-SNE-based feature visualization were assessed after 20 epochs of training. The experiment employed 5,000 training, 500 validation, and 1,000 testing images.

### *A. Training and Validation Performance*

Figure 1 shows the proposed CAXNet model's training and validation. The model improved gradually throughout training and attained 92.4\% validation accuracy at epoch 12. The validation accuracy was 92.2\% at the last epoch, suggesting model stability following convergence.

Validation performance fluctuated for early epochs. This is frequent in deep CNN training with data augmentation and a learning rate scheduler. After the training phase, validation accuracy stabilized at 90\% or above for most epochs. The model learnt discriminative feature representations from chest X-ray pictures through persistent training loss reduction.

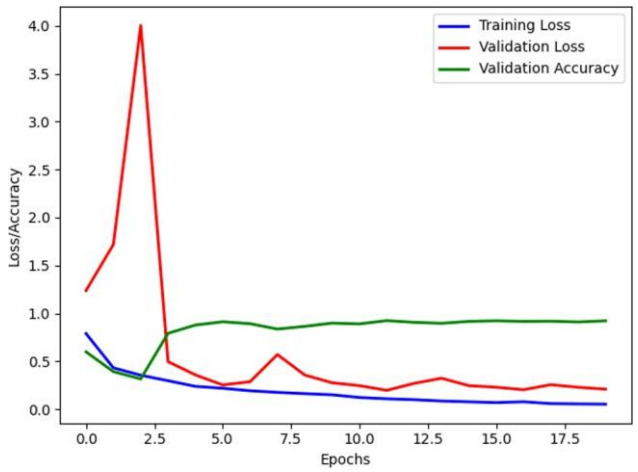


Fig. 1. Training and validation loss alongside validation accuracy over 20 epochs.

### *B. Overall Classification Performance*

The overall test performance of the proposed CAXNet model is shown in Table VI. The model categorized 920 of 1,000 chest X-ray pictures, achieving 92.00% accuracy.

**TABLE VI**
**OVERALL CLASSIFICATION PERFORMANCE OF CAXNET ON THE TEST SET**

| Metric | Value |
|---|---|
| Accuracy | 92.00% |
| Weighted Precision | 92.15% |
| Weighted Recall | 92.00% |
| Weighted F1-score | 92.05% |
| Matthews Correlation Coefficient | 0.8852 |
| Weighted ROC-AUC | 0.9329 |

The weighted precision, recall, and F1-score were 92.15%, 92.00%, and 92.05%. This shows that the model balanced categorization across the four target classes. A significant connection between anticipated and actual class labels is shown by the MCC value of 0.8852. MCC is useful for measuring classification performance when class distributions are not balanced since it incorporates all confusion matrix components.

The model's weighted ROC-AUC of 0.9329 suggests acceptable class-separation. These results demonstrate that CAXNet learns to differentiate COVID-19, Lung Opacity, Normal, and Viral Pneumonia chest X-ray images.

### *C. Class-wise Performance and Confusion Matrix Analysis*

The confusion matrix of the proposed model is shown in Figure 2. It provides a detailed view of the correctly and incorrectly classified samples for each class.

In the confusion matrix, CAXNet accurately categorized 240 of 250 COVID-19 images and 86 of 88 Viral Pneumonia images. This shows strong categorization for these two groups. The model properly categorized 385 of 425 Lung Opacity images. The Normal class categorized 209 of 237 images correctly.

Errors were mostly between Lung Opacity and Normal classes. 37 Lung Opacity and 22 Normal images were misclassified. When opacity-related defects are mild, lung opacity and normal radiographs may have comparable patterns. Overlap can make the model's decision boundary between these two groups harder.

The class-wise performance values are summarized in Table VII.

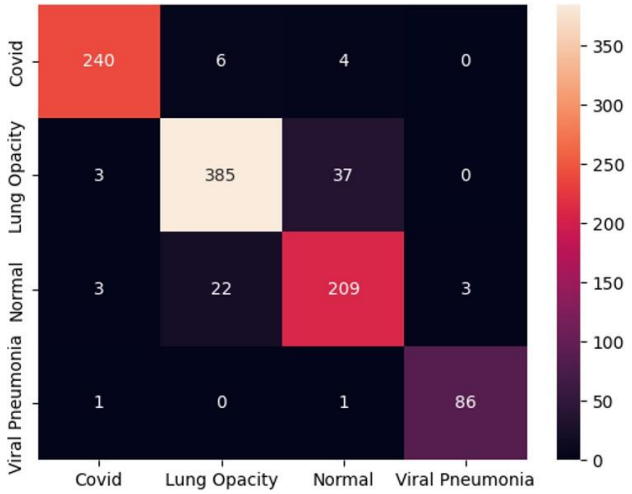


Fig. 2. Confusion matrix obtained by CAXNet on the test set.

The class-wise performance values are summarized in Table VII.

**TABLE VII**
**CLASS-WISE PERFORMANCE OF CAXNET ON THE TEST SET**

| Class | Correct / Total | Precision | Recall | F1-score |
|---|---|---|---|---|
| COVID-19 | 240 / 250 | 97.17% | 96.00% | 96.58% |
| Lung Opacity | 385 / 425 | 93.22% | 90.59% | 91.89% |
| Normal | 209 / 237 | 83.27% | 88.19% | 85.66% |
| Viral Pneumonia | 86 / 88 | 96.63% | 97.73% | 97.18% |

Viral Pneumonia had the greatest recall among all the four classes, at 97.73%, followed by COVID-19 at 96.00%. The proposed model accurately identified these two illness types. Because some Lung Opacity samples were predicted as Normal, the Normal class had the lowest precision of 83.27%. Confusion matrix also shows this observation.

The results imply that CAXNet learned robust discriminative representations using Coordinate Attention-enhanced ResNeXt feature extraction and multi-level feature fusion. Lung Opacity and Normal are confused, suggesting that region-specific analysis or lung segmentation might enhance classification performance in the future.

### *D. t-SNE-Based Qualitative Feature-Space Analysis*

Using deep features from CAXNet's penultimate layer, an t-SNE visualization was created to test discrimination of the learnt representations. Global average pooling produced 768-dimensional feature vectors that were projected into two dimensions.

The t-SNE plot shows how the model arranges test samples in learnt feature space qualitatively. The image indicates that samples from various classes create distinct zones. Strong class-wise recall and F1-score support clearer separation in COVID-19 and Viral Pneumonia samples. It appears that Lung Opacity and Normal samples overlap, validating the confusion matrix's claim that these two groups are harder to distinguish.

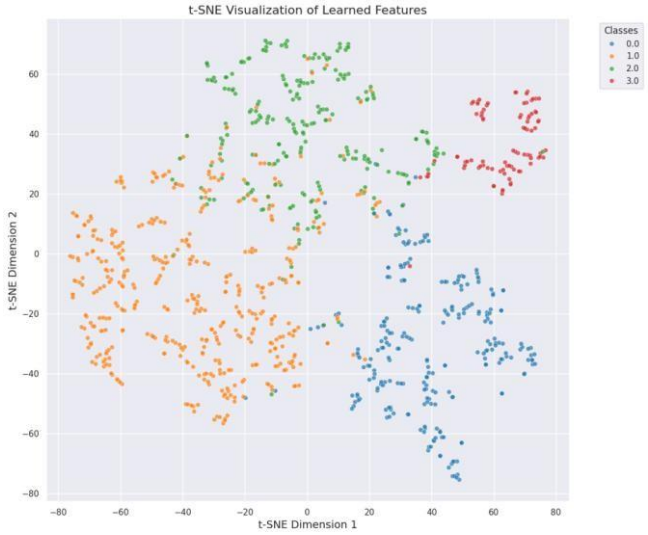


Fig. 3. t-SNE visualization of the learned features, showing class separability in the latent space.

The quantitative findings are strengthened by this qualitative examination. Chest X-ray pictures may teach CAXNet class-specific representations because to the substantially separable feature clusters. The partial overlap between Lung Opacity and Normal shows their visual similarities and explains the experiment's significant misclassification trend.

CAXNet performed well for multiclass COVID-19 chest X-ray classification with 92.00% test accuracy, 92.05% weighted F1-score, 0.8852 MCC, and 0.9329 weighted ROC-AUC. Using a modified ResNeXt backbone, Coordinate Attention-based feature refinement, and multi-level feature fusion, chest X-ray classification representations are successful. The confusion between Lung Opacity and Normal patients was the primary limitation, however COVID-19 and Viral Pneumonia performed well.

## VI. CONCLUSION

This paper introduces CAXNet, a Coordinate Attention-enhanced ResNeXt architecture for multiclass COVID-19 chest X-ray classification. The model alters the ResNeXt50-32×4d backbone by preserving just early and intermediate phases. After the retained residual stages, coordinate Attention modules augment spatially informative feature maps, and a multi-level feature fusion technique combines attention-enhanced low- and intermediate-level representations before classification.

The model was tested on a four-class chest X-ray classification task involving COVID-19, lung opacity, normal, and viral pneumonia. CAXNet has an overall test accuracy of 92.00%, weighted precision of 92.15%, recall of 92.00%, F1-score of 92.05%, MCC of 0.8852, and ROC-AUC of 0.9329. Results show that the suggested architecture can learn radiographic representations effectively and discriminatively.

The confusion matrix analysis indicated that the model accurately classified 240 of 250 COVID-19 images and 86 of 88 Viral Pneumonia images. The biggest categorization issue was Lung Opacity and Normal patients, where radiographic overlap caused misclassification. T-SNE showed moderately separable feature clusters for most classes and partial overlap between Lung Opacity and Normal samples, supporting this conclusion.

The results show that COVID-19 chest X-ray classification works using a compact ResNeXt-based feature extractor, Coordinate Attention-based feature refining, and multi-level feature fusion. Stratified data splitting, bigger and more diverse datasets, lung-region segmentation, external validation, and explainability approaches like Grad-CAM to further evaluate the model's decision-making areas can increase CAXNet's performance.

## REFERENCES

[1] E. F. Ohata, G. M. Bezerra, J. V. S. Chagas, A. V. L. Neto, A. B. Albuquerque, V. H. C. de Albuquerque, and P. P. R. Filho, “Automatic detection of COVID-19 infection using chest X-ray images through transfer learning,” IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 1, pp. 239–248, 2021.

[2] A. I. Khan, J. L. Shah, and M. M. Bhat, “CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images,” Computer Methods and Programs in Biomedicine, vol. 196, p. 105581, 2020.

[3] T. Mahmud, M. A. Rahman, and S. A. Fattah, “CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization,” Computers in Biology and Medicine, vol. 122, p. 103869, 2020.

[4] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, “Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network,” Applied Intelligence, vol. 51, no. 2, pp. 854–864, 2021.

[5] R. Jain, M. Gupta, S. Taneja, and D. J. Hemanth, “Deep learning based detection and analysis of COVID-19 on chest X-ray images,” Applied Intelligence, vol. 51, pp. 1690–1700, 2021.

[6] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. B. A. Kashem, M. T. Islam, S. Al Maadeed, S. M. Zughaier, M. S. Khan, and M. E. H. Chowdhury, “Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images,” Computers in Biology and Medicine, vol. 132, p. 104319, 2021.

[7] L. Kong and J. Cheng, “Classification and detection of COVID-19 X-ray images based on DenseNet and VGG16 feature fusion,” Biomedical Signal Processing and Control, vol. 77, p. 103772, 2022.

[8] M. Bhandari, T. B. Shahi, B. Siku, and A. Neupane, “Explanatory classification of CXR images into COVID-19, pneumonia and tuberculosis using deep learning and XAI,” Computers in Biology and Medicine, vol. 150, p. 106156, 2022.

[9] T. D. Pham, “Classification of COVID-19 chest X-rays with deep learning: New models or fine tuning?” *Health Information Science and Systems*, vol. 9, Art. no. 2, 2021.

[10] B. Oltu, S. Guney, S. E. Yuksel, and B. Dengiz, “Automated classification of chest X-rays: A deep learning approach with attention mechanisms,” BMC Medical Imaging, vol. 25, p. 71, 2025.