Ensemble fusion model for improved lung abnormality classification: Leveraging pre-trained models

Suresh Kumar Samarla [a,b,*], Maragathavalli P [a]

[a] Information Technology, Puducherry Technological University, Puducherry, India
[b] CSE Department, SRKR Engineering College, AndhraPradesh, India

ARTICLE INFO

ABSTRACT

Lung abnormalities pose significant health concerns, underscoring the need for swift and accurate diagnoses to facilitate timely medical intervention. This study introduces a novel methodology for the sub-classification of lung abnormalities within chest X-rays captured via smartphones. An accurate and timely diagnosis of lung abnormalities is essential for the successful implementation of appropriate therapy. In this paper, we propose a novel approach using a Convolutional neural network (CNN) with three maximum pooling layers and early fusion for sub-classifying lung abnormalities from chest Xrays. Based on the kind of abnormality, the CheXpert dataset is divided into 13 sub-classes, each of which is trained using a different sub-model. An early fusion procedure is then used to integrate the outputs of the sub-model.

- 3M-CNN (Method 1): We employed a Convolutional Neural Network (CNN) with three max pooling layers and an early fusion strategy to train dedicated sub-models for each of the 13 distinct sub-classes of lung abnormalities using the CheXpert dataset.
- Ensemble Model (Method 2): Our 'Ensemble model' integrated the outputs of the trained sub-models, providing a powerful approach for the sub-classification of lung abnormalities.
- Exceptional Accuracy: Our '3M-CNN' and 'fused model' achieved an accuracy of 98.79%, surpassing established methodologies, which is beneficial in resource-constrained environments embracing smartphone-based imaging.

Graphical abstract

Machine learning procedures have transformed the landscape of medical imaging, providing precise diagnostic tools for identifying various lung abnormalities such as pneumonia, COVID-19, tuberculosis, and more. These advanced techniques harness complex algorithms and extensive datasets to detect patterns and irregularities in medical images, empowering radiologists to render more accurate diagnoses. Among these methods, Convolutional Neural Networks (CNNs) are frequently employed, utilizing artificial neural networks to autonomously discern structures from raw or under-exposed images [8,9,10,11].

We introduce a unique strategy for sub-classifying lung abnormalities from smart-phone-captured photos using a modified Convolutional Neural Network (3M-CNN) with three max-pooling layers and early fusion. The dataset is divided into 13 subclasses based on the type of abnormality, with each having its dedicated sub-model (as depicted in Fig. 1). This approach aims to enhance the precision and effectiveness of lung abnormality detection and classification.

---

* Corresponding author at: Information Technology, Puducherry Technological University, Puducherry, India.
E-mail addresses: Samarla.sureshkumar@pec.edu, ssk@srkrec.ac.in (S.K. Samarla).
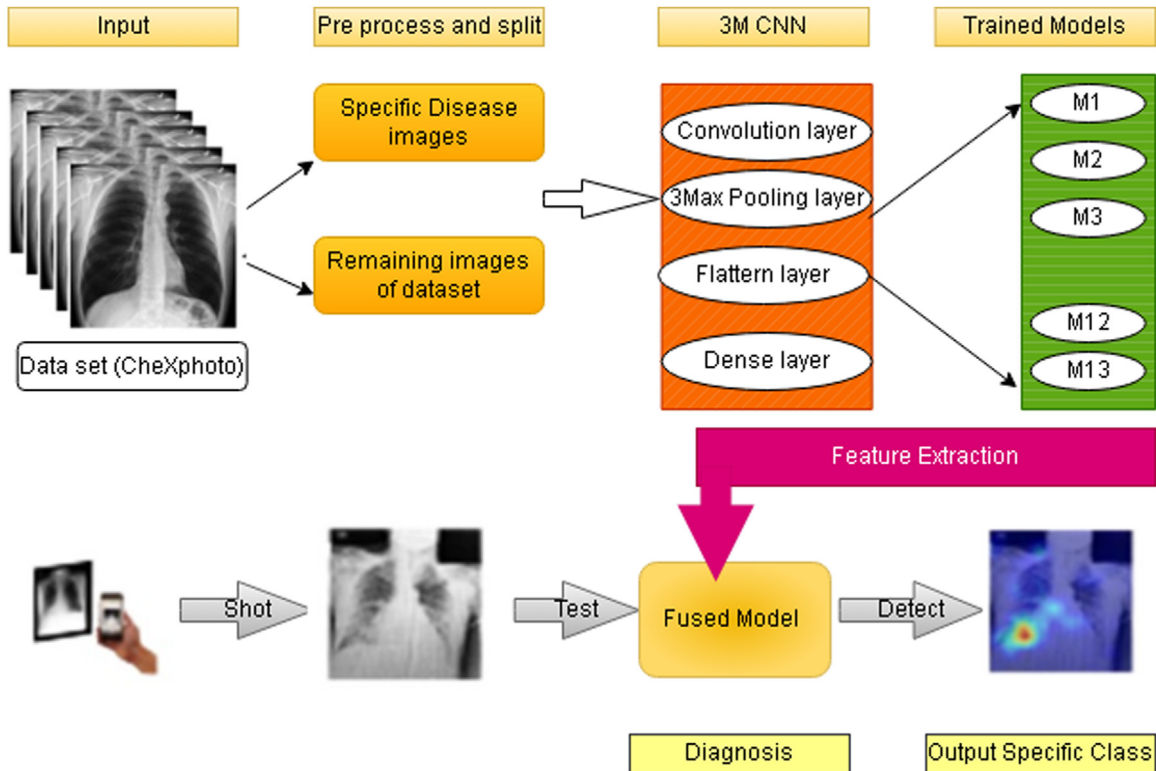
Fig. 1. Architecture of the proposed system.

Algorithm: Enhanced Lung Abnormality Classification
Step 1: Create Base Model

Input: Raw Data (Smartphone-Captured Photos)
Output: Base CNN Model (3M-CNN) for Binary Classification

Step 2: Train and Evaluate

For each Sub-Class in Dataset:

    - Train the Base Model for the Specific Sub-Class
       Evaluate the Trained Model's Performance

Step 3: Fusion of Fine-Grained Classification

Load Base Model and Sub-Class Models
Set all models to Evaluation Mode

Step 4: Build Fused Model

Create a New Model by Integrating Base Model with Sub-Class Models
Adapt Forward Function to Process Input via Base Model, Select Appropriate Sub-Class Model

Step 5: Evaluation

Evaluate Fused Model's Performance on Test Dataset
Measure Accuracy in Classifying Across All Sub-Classes

End Algorithm

Dataset

The Dataset used was ChestXpert. Which is a collection of chest Xrays divided into Train set and valid set. The Train Dataset is a comprehensive collection of medical images categorized into various path types, including digital, photographic, nokia, and iPhone images, with each type having around 10,507 to 1000 instances. Gender-wise, the dataset predominantly consists of 20,549 male and 11,972 female cases. The dataset encompasses a wide range of medical findings and conditions, with Lung Opacity being the most prevalent at 16,024 cases, followed by Support Devices at 17,255. Other notable conditions include Edema, Pleural Effusion, and Atelectasis, all contributing to the dataset's diversity and complexity. And valid Dataset is a smaller subset derived from the Dataset, containing only 234 instances with digital, photographic, and oneplus path types. Gender distribution is less skewed, with 384 males and 318 females. The findings and conditions are also present but in reduced numbers compared to the Train Dataset [1,2].

Specifications table

Subject area:	Computer Science
More specific subject area:	Deep learning
Name of your method:	Modified CNN (3M-CNN), Ensemble of Pretrained Models(Fused Model)
Name and reference of original method:	The proposed method consists of numerous techniques discussed and cited within the method details of this article.
Resource availability:	Hardware: A fast multi-core CPU or one or more powerful GPUs.
	Software: Deep Learning Framework like TensorFlow, Keras, PyTorch, or a similar library and Data Science Libraries such as NumPy, scikit-learn, and matplotlib for data manipulation, evaluation, and visualization.
	Dataset: CheXpert Dataset(https://stanfordmlgroup.github.io/competitions/chexpert/)

Method details

Method1: modified CNN (3M-CNN)

We conducted an exploration involving Convolutional Neural Network (CNN) architectures, experimenting with various configurations. This included the variation of max pool layers at the outset, leading to the identification of an architecture featuring 3 max pooling layers followed by a flatten layer and a Dense layer. To optimize the model, we utilized the Adam optimizer, employing default β-parameters ($\beta1 = 0.9$, $\beta2 = 0.999$), and a fixed learning rate of 0.001 throughout the training process. The basic structure of modified CNN is shown in Fig. 2 and A comprehensive overview of the model is provided in the subsequent section

Early fusion

Early fusion, also known as feature-level fusion, is a machine learning technique employed to integrate features from multiple modalities or sources at an early stage of processing before entering a learning algorithm. This approach is particularly relevant in medical image classification, such as X-ray analysis, where the fusion of diverse features enhances predictive accuracy.

In the realm of medical imaging, multiple modalities may encompass various scans (e.g., X-rays, CT scans, MRIs) or different types of features derived from the same images (e.g., texture, shape, intensity features).

Key Steps in Early Fusion for Image Classification:

1. Feature extraction: Start by taking out distinguishing features from every source or modality. The process include extracting texture features and shape features,that are necessary for proper classification of X-ray .
2. Normalization: It is crucial to standardize the scale of features extracted from different modalities to ensure consistency. Normalization aligns the scales, addressing variations that may arise from the inherent differences in modality-specific feature representations
3. Combination: Merges the features from different modalities into a cohesive feature vector. This integration is achieved through simple concatenation and averaging
   # Sample code for feature combination (assuming 'features_modality1' and 'features_modality2' are extracted features)
   combined_features = np.concatenate((features_modality1, features_modality2), axis=1)
4. Classification: Direct the combined feature vector into a classification algorithm.
   # Sample code for classification (using a hypothetical classifier 'model')
   model.fit(combined_features, labels)
5. Model Fusion: Load pre-trained models for each sub-classification, such as 'Abnormality,' 'Edema,' 'Cardiomegaly,' etc. These models are then used collectively to create a fused model for comprehensive image classification.
   # Sample code for loading pre-trained models
   models = {}
   model_names = [list of model names]
   for model_name in model_names:
   model = load_model(f'{model_name}.h5', compile=False)
   models[model_name] = model
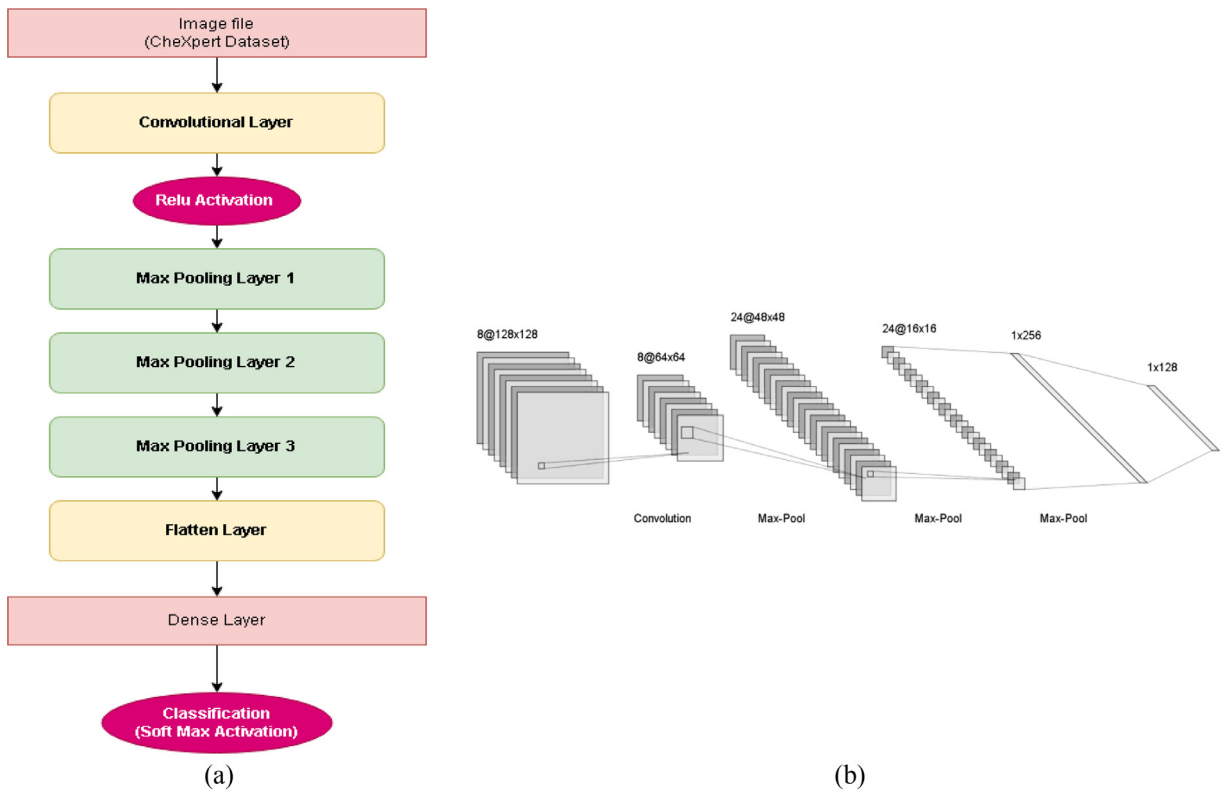
Fig. 2. Modified CNN Architectural view (Fig. 2.a (3M-CNN), b: Neural Network view)The "3M-CNN" (shown in Fig. 2.a and b) or Convolutional Neural Network with three Max Pooling layers, is a simplified and fundamental image categorization architecture. Despite its simplicity in comparison to more complex models such as MobileNet, GoogleNet, InceptionNet, and ResNet, this CNN setup has several advantages, particularly in terms of architecture and resource requirements.

6. Image Classification using Fused Model: The code iterates through each loaded model, makes predictions, and combines the results for a holistic classification.
   # Sample code for image classification using the fused model
   conditions, percentages = classify_image(image_path, models)

Summary of the Modified CNN(3M-CNN)
3M-CNN- Algorithm:
Let X be the input tensor of shape (batch_size, 224, 224, 3).
Convolutional layer:

- Input shape: (batch_size, 224, 224, 3)
- Filter shape: (3, 3, 3, 64)
- Padding: 'same'
- Strides: (2, 2)
- Activation: ReLU
- Output shape: (batch_size, 112, 112, 64)

Output tensor after convolution:
$Z\_1 = ReLU(X * W\_1 + b\_1)$
Max-pooling layers:

- Pool size: (2, 2), Strides: (2, 2), Output shape: (batch_size, 56, 56, 64)
- Output tensor after max pooling: $Z\_2 = max\_pool(Z\_1)$
- Pool size: (2, 2), Strides: (2, 2), Output shape: (batch_size, 28, 28, 64)
- Output tensor after max pooling: $Z\_3 = max\_pool(Z\_2)$
- Pool size: (2, 2), Strides: (2, 2), Output shape: (batch_size, 14, 14, 64)
- Output tensor after max pooling: $Z\_4 = max\_pool(Z\_3)$

Flatten layer:

  - Input shape: (batch_size, 14, 14, 64)
  - Output shape: (batch_size, 12,544)

Output tensor after flattening:
Z_5 = flatten(Z_4)
Dense layer:

  - Input shape: (batch_size, 12,544)
  - Units: 2
  - Activation: Softmax
  - Output shape: (batch_size, 2)

Output tensor after the dense layer:
Z_6 = Softmax(Z_5 * W_6 + b_6)
Model Compilation:

  • Optimizer: Adam
  • Loss function: Categorical cross-entropy
  • Evaluation metric: Accuracy

Summary of the Updated CNN: Input -> Conv -> MaxPool -> MaxPool -> MaxPool -> Flatten -> Dense -> Output
End Algorithm

The 3M-CNN has the following advantages over advanced CNN architectures:

  • **Simplicity and resource conservation:** The simplicity of the 3M-CNN's architecture makes it less computationally intensive than more sophisticated networks like InceptionNet or ResNet. It requires fewer parameters, which makes it more memory-efficient and faster to train, which is useful for applications with limited computational resources.
  • **Ease of Implementation:** Its basic structure makes it very straightforward to construct and understand, which might be advantageous for researchers or developers new to deep learning or seeking a simpler model to work with.
  • Faster Inference: The 3M-CNN's lower complexity may result in faster inference times, making it appropriate for real-time applications requiring rapid decision-making based on picture classification.
  • **Reduced Overfitting:** Because the 3M-CNN has fewer layers and parameters, it may be less prone to overfitting, especially when the dataset is small, making it more suitable for smaller datasets.
  • **Balanced Performance:** While more complex networks excel at specific jobs or have been fine-tuned for specific sorts of data, the 3M-CNN's basic design may perform more consistently across a wider range of data.

Method2: Ensemble of Pretrained Models
The method involves classifying lung abnormalities using pre-trained models, as shown in Fig. 3. The algorithm details the steps from data loading and model initialization to the image classification process and output results.
Algorithm: Lung Abnormality Classification

1. Data Loading and Model Initialization:
      • Load the dataset: Load the dataset containing images from the specified source.
      • Preprocessing: Prepare the images for analysis by standardizing their size (e.g., resizing to 224×224 pixels).
      • Initialize pre-trained models: Pre-load and initialize pre-trained models dedicated to various lung Abnormality classifications. These models, such as 'Abnormality.h5', 'Edema.h5', 'Cardiomegaly.h5', etc., are specifically designed for detecting different types of lung abnormalities.
2. Image Classification Process:
      • 'classify_image' function: This function processes each image from the dataset for Abnormality classification using the pre-loaded models.
      • For each image in the dataset:
      • Read and preprocess the image: Standardize the image sizes, converting them to a consistent format for analysis, often resized to a specified dimension like 224×224 pixels.
      • Iterate through pre-trained models: Use the pre-loaded models to predict abnormality conditions for each processed image.
      • Categorize abnormality conditions: Analyze prediction percentages generated by each model and apply specific thresholds to categorize the Abnormalities detected by each model.
3. Output the Result:
      • This stage likely involves organizing the detected Abnormalities, their respective conditions, and associated prediction percentages into a structured format, enabling further analysis or reporting. This output are exported for subsequent evaluation or reporting purposes.
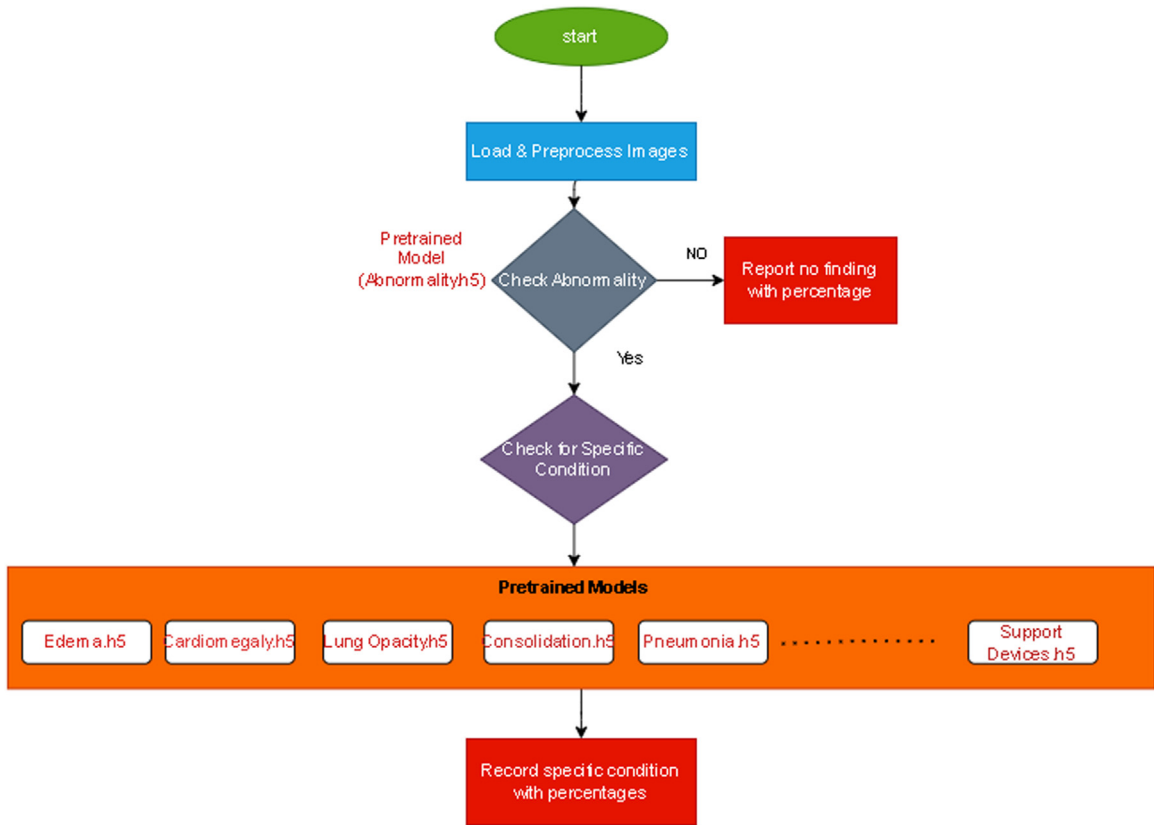
End Algorithm

Fig. 3. Ensemble of Pretrained Models.

Initially, we employed a standard pre-trained model trained on ImageNet, leveraging its capability to recognize fundamental features within the input images. This model was combined with a modified CNN focused on chest X-rays (CXR) to predict abnormalities. Subsequently, 14 distinct models were trained to recognize each sub-class. These pre-trained models were amalgamated to precisely identify the sub-class corresponding to a given input. After training the sub-models, we used early fusion to combine the predictions from each sub-model and the base model. The base model consisted of the same architecture as the sub-models, but was trained on the entire dataset without dividing it into sub-classes. We combined the predictions from each model using a weighted average approach, with weights determined by the performance of each sub-model on a validation set.

Validation of the ensemble method

The proposed Ensemble Method's performance is thoroughly evaluated in relation to individual models. For an extensive assessment, two important tables are provided. One (Table 2) "Result Comparison with the Existing System" and other (Table 3) "Performance Analysis of Individual Models".A comparison table (Table 2) shows the accuracy percentages for different classes across the ensemble method and individual models. These percentages represent the model's capability to correctly classify a specific class and a comprehensive review of the performance measures for each model across various kinds of lung conditions is shown in Table 3. Accuracy, precision, recall, F1-score, and support for every class. This table provides a starting point to determine the effectiveness of the Ensemble Method.

Key Observations:

* The ensemble method, as represented by the fused model, demonstrates promising performance across most medical conditions, often outperforming individual models and references.
* Conditions like Pneumothorax and Lung Lesion showcase notably high accuracy in the ensemble method compared to standalone models.
* However, some conditions such as Enlarged Cardiomediastinum and Fracture present challenges, showing lower accuracy across all models, which could require further investigation or different methodologies.

Overall, the ensemble method, as indicated by the fused model, displays robustness and improved performance in classifying lung Abnormalities, offering potential enhancements in accuracy compared to individual models and references. Further detailed analysis and comparative studies might help refine the ensemble approach for specific medical conditions requiring additional attention.

Table 2

Result comparison with the Existing System.

	[3]	[4]	[5]	[6]	[7]	Basic-CNN	Ensemble model
Cardiomegaly	81.00	85.60	88.3	85.40	81.19	70.31	**85.52**
Edema	80.50	80.60	83.5	93.90	72.15	75.90	**83.88**
Consolidation	70.30	71.10	74.5	86.50	**92.79**	62.38	**90.19**
Pneumonia	65.80	68.40	73.1	74.60	**79.18**	76.71	**86.23**
Atelectasis	70.00	73.30	76.7	84.50	82.88	68.29	**88.69**
Pneumothorax	79.90	80.50	84.6	95.60	89.96	58.70	**87.15**
Pleural Effusion	75.90	80.60	82.8	93.80	62.75	51.00	**73.31**
Pleural Other	68.40	72.40	76.1	96.50	**98.27**	85.71	**88.83**
Support Devices	–	–	–	94.60	61.40	54.86	**94.01**
Enlarged Cardiomediastinum	–	–	–	61.60	**68.67**	59.29	**66.24**
Lung Opacity	–	–	–	92.00	55.53	50.12	**92.77**
Lung Lesion	–	–	–	88.00	**95.68**	84.12	**94.78**
Fracture	–	–	–	0.00	**95.91**	70.97	**94.11**
Abnormality	–	–	–	0.889	**94.18**	90.17	**98.79**

Table 3

Performance analysis of individual models.

Class	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Support
Cardiomegaly	85.52	80.4	100	90.57	209
Edema	83.88	80.84	91	75.59	271
Consolidation	90.19	91.85	93.64	88.68	546
Pneumonia	86.23	76.45	72.68	82.65	364
Atelectasis	88.69	96.82	90.17	68.73	570
Pneumothorax	87.15	78.5	96.2	81.62	780
Pleural Effusion	73.31	51.02	48.5	55.33	330
Pleural Other	88.83	85.17	72.17	76	496
Support Devices	94.01	86.54	84.98	82	428
Enlarged Cardiomediastinum	66.24	59.29	62.8	67.9	710
Lung Opacity	92.77	86	92	96	364
Lung Lesion	94.78	84.12	96.24	98.7	540
Fracture	94.11	97.07	92.48	86.24	460
Abnormality	98.79	90	89.72	95	710

This validation process showcases the efficacy of ensemble methods in improving classification accuracy for lung Abnormalities, highlighting the potential for enhanced diagnostic precision and effectiveness in medical imaging analysis.

Ethics statements

The authors do not have permission to share data

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Suresh Kumar Samarla: Conceptualization, Methodology, Data curation, Formal analysis, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Maragathavalli P:** Supervision, Formal analysis, Investigation, Writing – review & editing.

Data availability

The authors do not have permission to share data.

Acknowledgments

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.mex.2024.102640

References

[1] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Ilcus, C. Chute, H. Marklund, B. Haghgoo, R. Ball, K. Shpanskaya, et al., Chexpert: a large chest radiograph dataset with uncertainty labels and expert comparison, in: Proceedings of the AAAI Conference on Artificial Intelligence, 33, 2019, pp. 590–597. pages.

[2] Stanford ML GroupChexpert: a Large Chest X-Ray Dataset and Competition, Jan 2019 https://stanfordmlgroup.github.io/competitions/chexpert/.

[3] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M Summers, arXiv preprint, 2017.

[4] L. Yao, E. Poblenz, D. Dagunts, B. Covington, D. Bernard, K. Lyman, arXiv preprint, 2017.

[5] S. Guendel, et al., arXiv preprint, 2018.

[6] H.H. Pham, T.T. Le, D.Q. Tran, D.T. Ngo, H.Q. Nguyen, Interpreting Chest X Rays via CNNs that Exploit Disease Dependencies and Uncertainty Labels, 2019 Preprint at https://arxiv.org/abs/1911.06475 .

[7] S. Kumar S, P.D.S.S. Lakshmi Kumaria, M.K.T.P ManikantaReddya, V.S.S.S Ramarajua, N. Pathak, Abnormality detection in smartphone-captured chest radiograph using multi-pretrained models, in: Proceedings of Data Analytics and Management, ICDAM, 2023 Volume 2ISBN978-981-99-6546-5, doi:10.1007/978-981-99-6547-2.

[8] L. Wang, Deep learning techniques to diagnose lung cancer, Cancer. (Basel) 14 (22) (2022) 5569 PMID: 36428662; PMCID: PMC9688236, doi:10.3390/cancers14225569.

[9] F.J.P. Montalbo, Truncating a densely connected convolutional neural network with partial layer freezing and feature fusion for diagnosing COVID-19 from chest X-rays, MethodsX 8 (2021) 101408 ISSN 2215-0161, doi:10.1016/j.mex.2021.101408.

[10] S. Bharati, P. Podder, M. Rubaiyat Hossain Mondal, , Hybrid deep learning for detecting lung diseases from X-ray images, Inform. Med. Unlock. 20 (2020) 100391 ISSN 2352-9148, doi:10.1016/j.imu.2020.100391.

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