LungNet: *Deep Learning Approach for Thoracic Disease Classification*

|  |  |  |
| --- | --- | --- |
| Aradhya Alva Rathnakar  *Dept. of Applied Data Science*  *San Jose State University*  San Jose, California, USA  aradhyaalva.rathnakar@sjsu.edu | Bhavan Kumar Basavaraju  *Dept. of Applied Data Science*  *San Jose State University*  San Jose, California, USA  bhavankumar.basavaraju@sjsu.edu | Mahamaya Panda  *Dept. of Applied Data Science*  *San Jose State University*  San Jose, California, USA  mahamaya.panda@sjsu.edu |
| Shashi Kumar Kadari Mallikarjuna  *Dept. of Applied Data Science*  *San Jose State University*  San Jose, California, USA  shashikumar.kadarimallikarjuna@sjsu.edu | Dr. Mohammad Masum  *Dept. of Applied Data Science*  *San Jose State University*  San Jose, California, USA  mohammad.masum@sjsu.edu |  |

***Abstract*—In the current day medical world, thoracic diseases such as pneumonia, effusion, and lung cancer pose significant diagnostic challenges due to the subtleties of radiological images and the high potential for human error in manual interpretation by radiologists. Traditional computer-aided diagnostic systems often fall short in capturing intricate patterns essential for accurate diagnosis, further exacerbated by the presence of false positives resulting from equipment artifacts or patient movement. This research introduces ‘LungNet’, an advanced deep-learning framework designed to revolutionize thoracic disease classification by enhancing both accuracy and efficiency in the diagnostic process. LungNet utilizes a two-stage deep learning approach to address these critical issues. In the first stage, the lung region is segmented from X-ray images using a U-Net architecture which acts as a crucial preprocessing step, isolating relevant anatomical structures and significantly reducing background noise and irrelevant features. The second stage involves an ensemble model combining DenseNet121 and ResNet50, which leverages the strengths of both architectures to achieve robust feature extraction and improved disease classification accuracy. This ensemble model is equipped with mechanisms like dropout layers and batch normalization to enhance noise tolerance and model generalization. The trained UNet model for lung segmentation resulted in an accuracy of 0.98, an F1 score of 0.97, and a Mean IoU of 0.96. The ensemble model for disease classification resulted in 85% accuracy. Overall, LungNet represents a significant advancement in medical imaging, offering a comprehensive solution that combines segmentation and classification into a unified diagnostic framework, thereby setting a new standard in healthcare delivery.**

***Keywords—LungNet, U-Net, Segmentation, ensemble, DenseNet121, ResNet50, disease classification***

# INTRODUCTION

LungNet is an innovative attempt at transforming the classification of thoracic diseases using deep learning techniques. Lung diseases such as pneumonia, effusion, and lung cancer are still difficult to diagnose accurately despite great strides made in medical imaging because radiologists must interpret them manually. This method takes a long time and is susceptible to mistakes made by humans who might miss minor irregularities. Moreover, traditional computer-aided diagnostic systems fail to detect intricate patterns in thoracic imaging data sufficiently. For instance, in X-ray images, false positives can occur during classification tasks because of noise from different origins like equipment artifacts or patient motion. Thus, there is a critical need for automated and more accurate analysis techniques to improve diagnostic outcomes and facilitate timely interventions.

To solve these problems, lungnet adopts a two-stage process. Firstly, the lung region is segmented from X-ray images with U-Net architecture which provides a specific area for further analysis. It serves as an important preprocessing step for disease classification by highlighting relevant anatomical structures, hence improving accuracy and efficiency. Additionally, the project proposes an ensemble model comprising DenseNet121 and ResNet50 for disease classification. By leveraging the strengths of each network and incorporating mechanisms for robust feature extraction and noise tolerance, such as dropout layers and batch normalization, the project aims to improve overall diagnostic accuracy and reliability. This innovative methodology fills a critical gap in the literature review by offering a comprehensive solution that combines segmentation and classification within a unified framework, ultimately transforming thoracic disease diagnosis and benefiting healthcare delivery.

# RELATED WORK

In recent years, rapid advancements have been made in using deep learning techniques for diverse medical image analysis tasks including segmentation, classification, and detection. Convolutional neural networks (CNNs) in particular, have attained state-of-the-art performance across modalities such as X-ray, CT, MRI, and histopathology images when sufficient labeled training data is available. For liver disease classification on ultrasound, Paul & Ramkumar [1] compared EfficientNet and SegNet, finding SegNet achieved a higher accuracy of 91% in distinguishing normal and fatty liver conditions. In digital pathology for colon cancer screening, Pandey et al. [2] adopted a UNet approach that compressed nucleus segmentation down to a single mask image per input. This attained 96.8% validation accuracy for cell detection, outperforming baselines for class imbalance problems. Manikandan et al. [3] evaluated multiple CNNs for pneumonia diagnosis on chest X-rays, determining that DenseNet201 delivered the highest sensitivity and specificity of 98.5% by focusing training on hard negative mining to overcome data imbalance. More broadly, for chest X-ray-based screening, Rajpurkar et al. [4] developed CheXNet, a 121-layer CNN-matched radiologist-level ROC curve performance for pneumonia detection across over 100,000 images. Wang et al. [12] released the large-scale ChestX-Ray8 dataset spanning 108,948 radiograph images across 8 common thoracic diseases, presenting a baseline classification and weakly-localized heatmap approach using ResNet-50 and log-sum-exp pooling to handle multi-label prediction. They note that further research is still required to achieve high-precision computer-aided diagnosis with such weakly labeled medical imaging data. Focusing model capacity via attention mechanisms on salient disease-specific image regions, Guan et al. [5] could further advance CNN performance for multi-label thoracic disease classification.

Medical imaging segmentation leverages automated delineation of anatomical structures like organs, bones, and lesions to assist in diagnosis, surgery planning, and treatment. Irshad et al. [9] demonstrated how multi-task learning through jointly predicting organ segmentations alongside boundary maps improves abdomen CT scan analysis accuracy over standard UNets. Multi-scale approaches have also shown benefits. Fang & Yan [10] learned the representations between partially labeled multi-organ CT datasets from different body areas via multi-scale CNNs. Trullo et al. [11] aptly combined fine-grained texture and global anatomical context modeling to enhance automated segmentation of thoracic organs at risk for radiotherapy planning. Finally, Dong et al. [15] presented an adversarial UNet-GAN system for multi-organ chest CT analysis, achieving high segmentation dice performance and sub-2mm surface distances between automated and manual organ delineations.

For peach crop pest management in agriculture, Bedi & Gole [6] leveraged unsupervised pre-training with convolutional autoencoders to improve the efficiency of bacterial spot disease CNN classifiers using leaf image data. Balavani et al. [14] talks about using the ResNet-50 model for plant disease classification, which diverges the existing use case for the ResNet-50 model by classifying 38 different plant disease categories. For clinical pneumonia screening, Reshan et al. [7] showed the viability of MobileNet fine-tuning on chest X-ray datasets lacking extensive labeled examples, reaching 94% test accuracy for detection without prohibitive medical imaging data requirements. According to WHO definitions of chest X-rays, Mahomed et al. [8] designed automated tools to detect pediatric pneumonia. Model-assisted screening could help standardize disease interpretation and monitoring by attaining performance approaching radiologist inter-observer agreements.

Gibson et al. [13] focused on 3D multi-organ segmentation, developing DenseVNet which extracts dense feature representations across varying anatomical scales to delineate abdominal CT scan structures accurately. Evaluation against prior state-of-the-art methods demonstrated that DenseVNet achieved significantly better Dice segmentation scores and boundary distance metrics across organs. Accurate automated abdomen segmentation could assist interventions like computer-aided endoscopy and surgery. Nonetheless, further clinical validation is still necessary to translate these tools into improved patient outcomes. Zhou et al. [18] leverage the ResNet-18 and DenseNet-121 model to diagnose intracranial hemorrhage in CT scanning where ResNet-18 provided an accuracy of 89.64% and DenseNet-121 provided 82.5% accuracy. In summary, given substantial labeled training data, deep CNNs have obtained state-of-the-art performance on diverse medical imaging tasks - particularly classification, detection, and segmentation problems. Challenges remain in perfecting accuracy, handling data constraints, and integrating algorithms smoothly into clinical practice to demonstrate patient benefit. As medical imaging datasets continue growing alongside computational power and algorithmic advances, deep learning is poised to transform medical computer-aided diagnosis and treatment planning.

# METHODOLOGY

The development of the proposed disease classification solution involves a two-stage approach. The first stage involves the training of the U-Net model for lung segmentation which is used to predict the mask for the lungs and is used for pre-processing the x-ray images that are used for disease classification. The second stage involves the usage of the pre-processed images which have the lung region segmented for disease classification using an ensemble model developed using ResNet 50 and DenseNet 121.

## Dataset Description and EDA

The data for training the U-Net model is sourced from the Kaggle repository that is published by Nikhil Pandey called Chest X-Ray Masks and Labels [16] which has the X-ray images along with the segmentation masks that are used for the UNet model training and testing. The dataset consists of 800 X-ray image and segmentation mask pairs as a part of the training set and 96 X-ray image and segmentation mask pairs as a part of the test set, constituting 5.4GB in size. The actual X-ray image and segmentation sets are well-labeled and stored in .png format. The segmentation mask images are labeled the same as the X-ray images followed by ‘\_mask’ to differentiate between the actual X-ray images and the mask images for ease of working with the data. Fig. 1. shows the sample X-ray image and segmentation mask set that is used for training the UNet model.

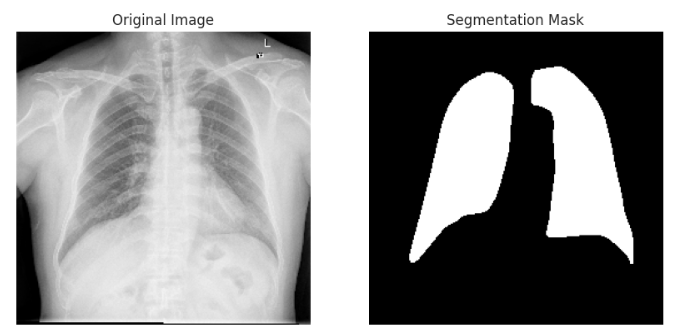
****

Fig. 1. Sample X-ray and lung segmentation mask.

The ChestX-ray-NIHCC dataset [17] is used for lung disease classification and it contains frontal view X-ray images of chests belonging to different patients along with the medical diagnosis. The dataset is contributed by the United States National Institutes of Health Clinical Center and consists of 112,120 frontal-view X-ray images of 30,805 unique patients with 14 base thoracic pathologies with a size over 45GB. The images are labeled with their disease findings such as No Finding, Edema, Consolidation, etc. The source contains a CSV file that has all the details related to the X-ray images, their diagnosis, and the bounding box location to point the disease in the X-ray to help with classification model training. The NIHCC website also provides a CSV file that has all the information about the patient diagnosis for all the X-rays available for classification purposes. The CSV file contains data about the X-ray images, diagnosed disease labels, patient gender, age, view position, image dimensions, and the follow-up visit number. There is also another CSV file containing the bounding box information on the location of the disease in the X-ray.   
Fig. 2. shows the sample X-ray image with the corresponding disease classification label.



Fig. 2. Sample chest X-ray for disease classification.

The patient diagnosis information is determined through EDA, and it describes the bounding box information of the disease location. Exploratory Data Analysis is performed to understand the various kinds of patient data that are being dealt with. The bar graph is plotted to Gender distribution in data where there are more than 60,000 X-rays for males and around 48,000 X-rays for females. Age distribution is analyzed to see in what range most patients fail to understand when lung-related diseases occur usually. The age distribution of the patients who got the X-ray is above 50 years of age. Fig. 3. delves into the distribution of view positions among patients who underwent X-ray examinations. It provides valuable insights into the preferred positioning of patients during radiographic procedures.

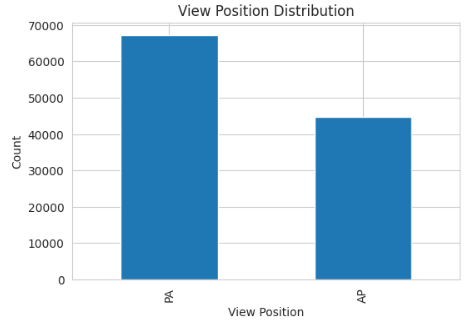


Fig. 3. View Position Distribution of Patients.

## Data Pre-processing

The data sourced from Kaggle for the lung segmentation task consists of well-labeled images enabling the ease of usage for the segmentation task. As an initial step, the raw X-ray images and the segmentation images are read simultaneously and undergo data pre-processing where the images are read in greyscale and the pixel values are normalized by dividing by 255. As a part of data preparation for creating train and validation sets, the source data, which is already split into train and validation sub-folders is read. The source images without the segmentation masks are dropped since they hold no value for training and validation. The raw X-ray image forms the X\_train and X\_val sets while the segmentation masks form the y\_train and y\_val sets.

Fig. 4. shows the top 15 most diagnosed lung diseases based on the source data which is essential to understand how common a disease is. It would also help in splitting the data for training and testing of the classification model ensuring there is a fair distribution of data for training and testing ensuring data stratification. Most of the X-rays result in no findings, followed by Infiltration, Atelectasis, etc.

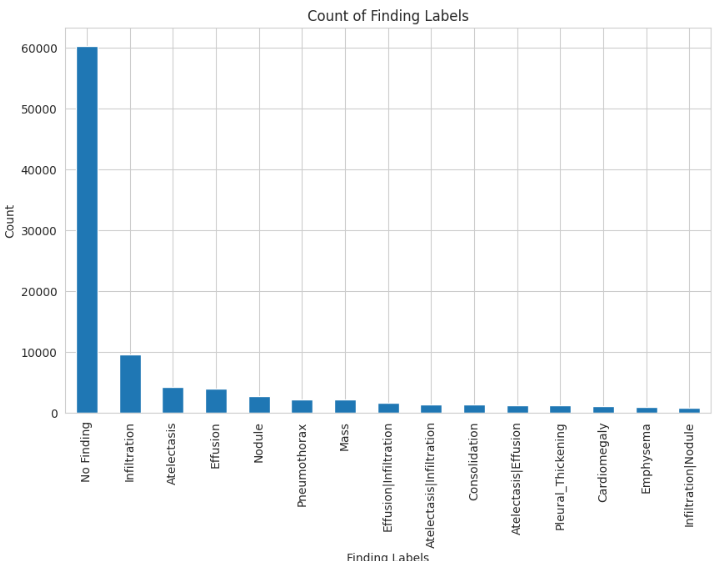


Fig. 4. Disease distribution.

Data preprocessing for lung segmentation also includes the usage of data augmentation techniques to increase the variety of data to improve the generalization of the segmentation model. The various augmentation techniques employed include Compose, HorizontalFlip, VerticalFlip, RandomRotate90, ShiftScaleRotate, RandomBrightnessContrast, and Resize. Fig. 5. shows a sample set of original and augmented images along with the mask.

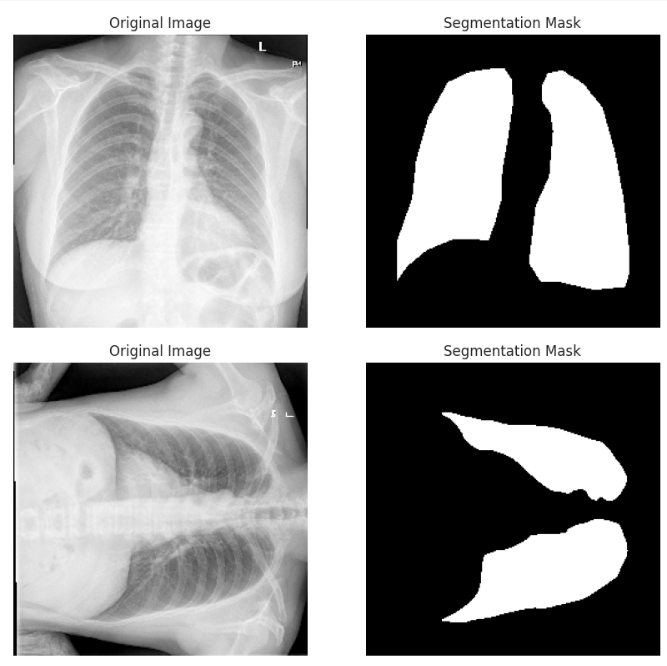


Fig. 5. Sample data after augmentation.

The prepared dataset for modeling includes 905 images for training and 227 images for validation with input image dimensions of 256 x 256 x 1. The UNet model is trained with the above data and is used for pre-processing the disease classification images where only the segmented lung image is considered for disease classification isolating the noise to improve the classification results.

The CSV file provided by NHICC is read and the corresponding X-ray images are pre-processed by performing lung segmentation using the UNet model which is trained before working on disease classification where the segmented lung region is the region of interest for disease classification. Fig. 6. shows the image after segmenting the source X-ray images for disease classification.

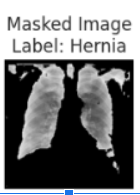


Fig. 6. Sample source and segmented data after pre-processing.

The pre-processed images for classification resulted in 31,395 images belonging to 10 unique base diseases with input X-ray image dimensions of 256 x 256 x 1. This pre-processed data is then split into an 80:10:10 ratio for training, validation, and test dataset for modeling the classification model.

# PROPOSED MODEL ARCHITECTURE

The proposed model architecture effectively harnesses the pre-processed data derived from structured directories. The initial phase involves training U-Net and SegNet models to perform lung segmentation, utilizing pairs of masked and original images. The dataset, meticulously arranged to segregate training, validation, and test sets, ensures that test data remains unseen during the training phase. Fig. 7. shows the detailed solution architecture for LungNet.

A diagram of a medical procedure

Description automatically generated

Fig. 7. Two stages of LungNet solution implementation.

## Segmentation Models

U-Net and SegNet architectures are structured with four convolutional layers (Conv2D) in the encoder, incorporating batch normalization, activation functions, and max pooling. A central layer bridges to the decoder, which mirrors the encoder’s structure but utilizes upsampling instead of pooling. The output layer employs a sigmoid activation function to classify pixels as foreground or background within the range of [0,1].

The U-Net model, distinguished by its skip connections that concatenate feature maps from each encoder layer to the corresponding decoder layer, preserves crucial information lost during down-sampling. This architectural nuance significantly enhances detail capture, making U-Net particularly effective for our dataset of approximately 1,000 image pairs.

In terms of optimization and tuning, firstly, U-Net’s performance, as measured by the mean Intersection over Union (mean IoU) score, suggests scope for improvement. To address this, the complexity is increased by adding additional convolutional layers to the encoder and decoder. The binary cross-entropy loss function was utilized to fine-tune foreground-background segmentation.

Further refinement was achieved through Random Search hyperparameter tuning, focusing on maximizing the mean IoU score. This method randomly samples combinations of learning rate, epochs, and batch size from a defined space, efficiently identifying optimal settings after several trials.

## Ensemble Classification Model

The segmentation model is used in the disease classification stage as a pre-processing step to segment the lung region, where the segmented images are paired with their corresponding labels. The ensemble approach is employed by combining ResNet-50 and DenseNet-121, leveraging transfer learning with pre-trained weights from the SwaV architecture [22], to enhance feature extraction capabilities. This setup adopts global average pooling to concatenate outputs from both models, effectively managing memory usage and mitigating overfitting.

The classification phase targets a subset of 10 labels from over 100 potential categories, using categorical cross-entropy for loss calculation and softmax for output activation. Strategic data stratification ensures balanced class representation, and additional dropout layers interspersed with dense ReLU activations further stabilize the model.

In summary, LungNet’s model architecture combines the detail-oriented capabilities of U-Net with the robust feature extraction and architectural diversity of an ensemble of ResNet-50 and DenseNet-121, tailored to efficiently handle both segmentation and classification tasks in thoracic disease diagnosis.

# EXPERIMENTAL SETUP

## Experimental Setup for U-Net

The experimental setup for the U-Net architecture is meticulously designed to achieve accurate lung segmentation from chest X-ray images. Below are the detailed procedures involved.

* + 1. *Data Preprocessing and Augmentation:* The training data comprises the chest X-ray images, and segmentation masks sourced from a structured directory. The original image and masks were processed to a uniform size of 256x256 pixels.

To enhance model generalization and prevent overfitting, several augmentation techniques were employed, including horizontal and vertical flips, 90-degree rotations, shifts, scales, and random brightness and contrast adjustments.

* + 1. *Model Architecture:* The U-Net model features symmetrical architecture with an encoder and decoder pathway. The encoder consists of repeated blocks of two convolutional layers (with 64, 128, 256, and 512 filters), each followed by batch normalization and ReLU activation, and a max pooling layer. The decoder mirrors the encoder structure but uses transposed convolutions for upsampling.

Critical for the U-Net architecture, skip connections are utilized to concatenate feature maps from the encoder to the corresponding layers in the decoder, preserving important features at different scales.

* + 1. *Training Setup:* The model was trained using binary cross-entropy as the loss function. Performance metrics included the Mean Intersection over Union (MeanIoU) to assess the accuracy of the segmentation.

An Adam optimizer with a dynamic learning rate was used, where the rate decreased exponentially after the initial 10 epochs to stabilize training progress.

* + 1. *Hyperparameter Tuning:* Hyperparameter tuning was conducted using Random Search to optimize learning rates among [0.001, 0.0001, 0.00001]. This strategy facilitated the exploration of the hyperparameter space more efficiently than grid search methods.

The model underwent several trials, each involving multiple executions, to robustly estimate the impact of hyperparameter choices on performance.

* + 1. *Training and Validation:* The dataset was split into training, validation, and test sets using an 80:10:10 ratio, ensuring sufficient data was available for model evaluation and to prevent information leakage.

The model was trained for 100 epochs with a batch size of 16, providing a balance between computational efficiency and model performance.

* + 1. *Performance Evaluation:* The model’s performance was evaluated on the validation and test sets. Metrics such as Mean IoU, F1 score, and accuracy were computed to quantify the model's effectiveness in segmenting lung regions from X-rays.

This structured approach underpins the robustness of the U-Net model in segmenting lung images and ensures reproducibility and reliability in medical imaging analysis.

## Experimental Setup for Ensembling Technique and Lung Segmentation

The experimental setup for combining lung segmentation using U-Net and disease classification using an ensemble of ResNet-50 and DenseNet-121 involves several stages from data preparation to model evaluation. Below are the detailed procedures involved.

* + 1. *Data Preparation and Preprocessing:* The dataset includes chest X-ray images and corresponding disease labels sourced from structured directories and CSV files containing diagnostic information.

Each X-ray image was resized to 256x256 pixels and converted to grayscale to standardize the input data format. The preprocessing also included normalization by dividing pixel values by 255.

* + 1. *Lung Segmentation:* The pre-trained U-Net model was loaded from 'unet\_model.h5' for segmenting lung regions from X-ray images.

Lung regions in the images were segmented by the U-Net model, followed by thresholding to convert the segmentation results into binary masks. Fig. 8. shows the output after segmentation of the original image.

A collage of x-ray images of lungs

Description automatically generated

Fig. 8. Output after segmentation of the original image.

* + 1. *Feature Extraction:* After lung segmentation, binary masks were applied to the original images to isolate lung regions. This step ensures that subsequent classification focuses on relevant areas, enhancing the accuracy and relevance of diagnostic predictions.
    2. *Ensemble Model for Disease Classification:* The ensemble model consisted of two base models, ResNet50 and DenseNet121, modified for grayscale input. Global average pooling was applied to each base model's output, which was then concatenated.

A dropout layer was added for regularization, followed by a dense layer with ReLU activation and a final dense layer with softmax activation to classify the images into one of ten disease categories.

* + 1. *Model Training and Validation:* The dataset was split into training and test sets with a ratio of 80:10:10 to evaluate the model's performance on unseen data.

The ensemble model was compiled with Adam optimizer and categorical cross-entropy loss function. The model was trained on the preprocessed images labeled with disease categories, using accuracy and recall as metrics.

* + 1. *Model Evaluation:* After training, the model's performance was evaluated on the test set, focusing on loss, accuracy, and recall assessing both the effectiveness and reliability of the predictions.

Exception handling was implemented during model fitting to ensure stability and diagnose potential issues in training.

* + 1. *Visualization and Interpretation:* Predictions were visually displayed by overlaying predicted disease labels on X-ray images. This step is crucial for practical diagnosis and allows medical professionals to examine the output obtained from the model’s prediction.

By integrating lung segmentation with ensemble classification, this experimental setup enhances the specificity of disease detection in thoracic X-rays. Fig. 9. shows the final prediction output for disease classification along with the segmented image.

A screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot

Description automatically generated A screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot of a screenshot

Description automatically generated

Fig. 9. Final prediction output along with segmented image.

The project's experimental setup and implementation can be found at <https://github.com/bhavanbk2/255>

## Hardware and Software Requirements

The development of LungNet is resource-intensive since it is an image-based application and requires higher GPU specifications for optimal and accelerated training without interruption. Jupyter Lab is used for the development of LungNet. The Nvidia P-100 GPU with 256 GB RAM is used for the training of UNet and ensemble (ResNet-50 + DenseNet-121) models.

Python is primarily used for the development of LungNet. Pandas 2.2.2 version of the library is used to perform data analysis. The Numpy 3.2 version of the library is used for dealing with data frames. Matplotlib 3.7 and Seaborn 0.11 versions are used to perform analysis and visualize results. Albumentations 1.4.7 version of the library is used to perform data augmentation to increase the generalization of the trained UNet model. Tensorflow version 2.16.1 and TF-Keras version 2.16.0 library is used for modeling UNet and the ensemble classification models. KerasTuner version 1.4.7 library is used for hyperparameter tuning of the UNet model. Scikit-learn library of version 1.4.2 is used to prepare the pre-processed data into train, validation, and test sets along with label encoding for disease classification.

# RESULT AND ANALYSIS

The experimental outcomes for the U-Net model and the Ensemble Model (ResNet 50 + DenseNet 121) elucidate their performance metrics through a comparative analysis with existing benchmarks in research. The ensemble model achieved a commendable test accuracy of 85.36%, which situates it competitively within the range reported by research benchmarks for similar architectures: 82.5% for DenseNet-121 and 89.64% for ResNet-18 models [18]. Notably, the model attained a recall of 85%, underscoring its capability to minimize false negatives, crucial in medical diagnostics, though it trails the higher recall of 93.3% reported in research [20].

In terms of segmentation accuracy, the U-Net model showcased a significant improvement in the Mean Intersection over Union (IoU) score of 0.96 after hyperparameter tuning, highlighting the effectiveness of tuning strategies. This post-tuning Mean IoU compares favorably against a research benchmark of 93.3% [20], demonstrating the model's precision in lung segmentation. Furthermore, the U-Net model achieved an F1 score of 0.9694 and an accuracy of 98.44%, both of which surpass the research benchmark scores of 73% and 95.4%, respectively [21]. These results emphasize the model's superior performance in accurately delineating lung contours from X-ray images, attributed to its robust architecture and training strategies.

1. Evaluation Metrics Comparison

| **Model​** | **Evaluation Metric** | **Proposal Results​** | **Comparative Benchmark Results** |
| --- | --- | --- | --- |
| ​  Ensemble Model​  ​  (ResNet 50 + DenseNet 121)​ | Test Accuracy​ | 0.8536​ | DenseNet-121: 0.825 [18] |
| ResNet-18: 0.8964 [18] |
| DenseNet-121: 0.835 [19] |
| Test Recall ​ | 85%​ | N/A |
| ​  UNet Model​ | Mean\_IoU score | 0.96​ | 93.3%​ [20] |
| F1 Score | 0.9694​ | 0.73​ [21] |
| Accuracy​ | 0.9844​ | 95.4%​ [21] |

In summary, the analysis indicates that while the Ensemble Model and U-Net Model align closely with, and in some metrics surpass, the performance of established benchmarks, opportunities for enhancement remain, particularly in improving the recall of the ensemble model. This could potentially be addressed by refining ensemble strategies or advancing image preprocessing techniques. These insights are instrumental for the further development of deep learning applications in radiological diagnostics, reinforcing the efficacy and potential of the proposed models.

# CONCLUSION

LungNet is developed based on a two-stage architecture, which involves lung segmentation followed by disease classification. The UNet model is trained for lung segmentation and is used to pre-process the X-ray images for disease classification. The UNet model developed, performed well with a Mean IOU of 0.96, accuracy of 0.98, and F1 score of 0.97. An ensemble model developed using ResNet 50 and DenseNet 121 that is used for the disease classification addresses challenges like vanishing gradient and overfitting. LungNet’s ensemble disease classification model demonstrates competitive performance achieving close to 85% accuracy. The unique approach of segmenting the organ of interest before performing disease classification holds the potential for transforming organ disease classification.

# FUTURE SCOPE

LungNet's impact on society is profound, as it offers improved diagnostic accuracy and timely interventions, reducing human error and optimizing healthcare resources. By automating the analysis of thoracic diseases, LungNet provides support to medical professionals, particularly in regions with limited access to specialized radiologists. However, the model faces limitations such as dependency on high-quality data, significant computational requirements, and challenges in clinical integration. Addressing these limitations through further refinement and ensuring ethical and regulatory compliance will enhance LungNet's potential to transform thoracic disease diagnosis and improve patient outcomes. Further fine-tuning can leverage learned features more effectively, particularly for rare or complex diseases. Incorporating advanced interpretability techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) and attention mechanisms will enhance model transparency and aid medical professionals in understanding diagnostic pathways, increasing trust in AI-driven systems. Seamless integration into healthcare systems, adherence to regulatory and ethical standards, and user-friendly interfaces are crucial for practical deployment. Additionally, employing active learning strategies could enhance model robustness by leveraging complex patterns in extensive medical image datasets, particularly those with 500+ diverse classes. This approach allows models to adapt to new data and emerging patterns, optimizing training processes and computational resources. By addressing these aspects, future initiatives can refine diagnostic accuracy, tackle practical deployment challenges, and foster broader acceptance and utilization of AI in medicine, potentially transforming patient outcomes through enhanced diagnostic capabilities.

##### ACKNOWLEDGEMENT

We gratefully acknowledge the contributions of all those who supported and facilitated the completion of this research project. Special thanks to Prof. Mohammad Masum for his guidance and supervision throughout the project. We also extend our gratitude to Teacher Assistant Vidushi Bhati for her prompt and continued support. We would also like to thank San Jose State University, Department of Applied Data Science for providing access to GPU which helped in the successful experimentation of the proposed project. Lastly, we acknowledge the invaluable contribution of the medical professionals and institutions who provided access to the necessary data and resources for this study.

##### REFERENCES

1. Paul, G., & Ramkumar, G. (2023, December). An Automated Classification and Segmentation of Fatty Liver Disease Using SegNet Model on Ultrasound Images. In *2023 International Conference on Emerging Research in Computational Science (ICERCS)* (pp. 1-6). IEEE.
2. Pandey, R., Lalchhanhima, R., & Singh, K. R. (2020, December). Nuclei Cell Semantic Segmentation Using Deep Learning Unet. In *2020 Advanced Communication Technologies and Signal Processing (ACTS)* (pp. 1-6). IEEE.
3. Manikandan, J., Shyni, S. A., Dhanalakshmi, R., Akshaya, S. V., & Dharshini, S. (2023, May). Segmentation and Detection of Pneumothorax using Deep Learning. In *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 468-473). IEEE.
4. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*.
5. Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., & Yang, Y. (2018). Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification. *arXiv preprint arXiv:1801.09927*.
6. Bedi, P., & Gole, P. (2021). Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network. *Artificial Intelligence in Agriculture*, *5*, 90-101.
7. Reshan, M. S. A., Gill, K. S., Anand, V., Gupta, S., Alshahrani, H., Sulaiman, A., & Shaikh, A. (2023, May). Detection of pneumonia from chest X-ray images utilizing mobilenet model. In *Healthcare* (Vol. 11, No. 11, p. 1561). MDPI.
8. Mahomed, N., van Ginneken, B., Philipsen, R. H., Melendez, J., Moore, D. P., Moodley, H., ... & Madhi, S. A. (2020). Computer-aided diagnosis for World Health Organization-defined chest radiograph primary-endpoint pneumonia in children. *Pediatric radiology*, *50*, 482-491.
9. Irshad, S., Gomes, D. P., & Kim, S. T. (2023). Improved abdominal multi-organ segmentation via 3D boundary-constrained deep neural networks. *IEEE Access*, *11*, 35097-35110.
10. Fang, X., & Yan, P. (2020). Multi-organ segmentation over partially labeled datasets with multi-scale feature abstraction. *IEEE Transactions on Medical Imaging*, *39*(11), 3619-3629.
11. Trullo, R., Petitjean, C., Ruan, S., Dubray, B., Nie, D., & Shen, D. (2017, April). Segmentation of organs at risk in thoracic CT images using a sharpmask architecture and conditional random fields. In *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)* (pp. 1003-1006). IEEE.
12. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2097-2106).
13. Gibson, E., Giganti, F., Hu, Y., Bonmati, E., Bandula, S., Gurusamy, K., ... & Barratt, D. C. (2018). Automatic multi-organ segmentation on abdominal CT with dense V-networks. *IEEE transactions on medical imaging*, *37*(8), 1822-1834.
14. Balavani, K., Sriram, D., Shankar, M. B., & Charan, D. S. (2023, May). An Optimized Plant Disease Classification System Based on Resnet-50 Architecture and Transfer Learning. In *2023 4th International Conference for Emerging Technology (INCET)* (pp. 1-5). IEEE.
15. Dong, X., Lei, Y., Wang, T., Thomas, M., Tang, L., Curran, W. J., ... & Yang, X. (2019). Automatic multiorgan segmentation in thorax CT images using U‐net‐GAN. *Medical physics*, *46*(5), 2157-2168.
16. *Chest xray masks and labels*. (2019, January 21). Kaggle. https://www.kaggle.com/datasets/nikhilpandey360/chest-xray-masks-and-labels
17. *Box*. (n.d.). https://nihcc.app.box.com/v/ChestXray-NIHCC
18. Zhou, Q., Zhu, W., Li, F., Yuan, M., Zheng, L., & Liu, X. (2022). Transfer learning of the ResNet-18 and DenseNet-121 model used to diagnose intracranial hemorrhage in CT scanning. *Current Pharmaceutical Design*, *28*(4), 287-295.
19. Liu, J., & Huang, Y. (2020). Comparison of different CNN models in tuberculosis detecting. *KSII Transactions on Internet and Information Systems (TIIS)*, *14*(8), 3519-3533.
20. Ferdinandus, F., Yuniarno, E. M., Purnama, I. K. E., & Purnomo, M. H. (2022, June). Covid-19 lung segmentation using U-Net CNN based on computed tomography image. In *2022 IEEE 9th International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)* (pp. 1-6). IEEE.
21. Saood, A., & Hatem, I. (2021). COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. *BMC Medical Imaging*, *21*, 1-10.
22. Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., & Joulin, A. (2020). Unsupervised learning of visual features by contrasting cluster assignments. *Advances in neural information processing systems*, *33*, 9912-99