

Report

by Aryan Ali

Submission date: 06-Dec-2023 09:50AM (UTC+0530)

Submission ID: 2249625505

File name: AryanAli_B20279_MTP1.pdf (2.38M)

Word count: 4444

Character count: 28496

Chest X-Ray Analysis

²
B.Tech Major Technical Project Report

to be submitted by

Aryan Ali

B20279

for the partial fulfillment of the degree

of

**BACHELOR OF TECHNOLOGY IN
DATA SCIENCE AND ENGINEERING**



SCHOOL OF COMPUTER AND ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY MANDI

KAMAND-175075, INDIA

NOVEMBER, 2023

Contents

1	Introduction	2
1.1	Objective	2
1.2	Motivation	3
1.3	Existing literature	3
1.4	Expected Outcome	4
2	Dataset and Data Preprocessing	5
2.1	Dataset (Initial findings)	5
2.2	Dataset Overview	5
2.3	Data Preprocessing	6
2.4	Data Splitting and Validation	7
3	System Model	8
3.1	U-Net	9
3.2	DeepLabV3+	10
3.3	Res U-Net	11
3.4	GCN (Global Convolution Network)	13
3.5	DenseNet-121	16
3.6	Bounding Box Generation	17
3.7	Weighted Binary Cross-Entropy	19
4	Results	21
5	Future Work	25

List of Tables

4.1 Result Comparisons for image segmentation	22
---	----

List of Figures

3.1	<i>Methodology adopted</i>	8
3.2	<i>Architecture of U-Net</i>	10
3.3	<i>Architecture of Deeplabv3+</i>	11
3.4	<i>Architecture of Res U-Net</i>	12
3.5	<i>Architecture of GCM and BRM</i>	14
3.6	<i>Architecture of Segmentation Model</i>	15
3.7	<i>Architecture of Global Convolution Network</i>	15
3.8	<i>Architecture of DenseNet Model</i>	17
3.9	<i>Bounding Box work flow</i>	17
3.10	<i>Class Activation Mapping</i>	18
3.11	<i>Gradient-weighted Class Activation Mapping</i>	19
4.1	¹² <i>Loss and dice coefficient values for UNet model</i>	21
4.2	¹² <i>Loss and dice coefficient values for ResUNet model</i>	22
4.3	<i>Loss and AUC values for GCN model</i>	22
4.4	<i>AUC score comparison for different diseases with existing works[1] [2]</i>	23
4.5	<i>Showing the final output with bounding box</i>	24
6.1	<i>Detailed timeline</i>	26

Chapter 1

Introduction

Chest X-rays (CXR) are vital diagnostic tools, offering a non-invasive means to assess the health of the chest, including the heart, lungs, and surrounding structures. Widely used in medical diagnostics, CXRs help identify and screen various lung conditions, such as pneumonia, tuberculosis, and lung cancer. They play a crucial role in emergency situations, providing rapid insights into injuries or trauma. With the ability to detect abnormalities and guide early intervention, CXRs contribute significantly to patient care, making them an indispensable component of modern healthcare for their efficiency in capturing detailed images of the chest's internal structures.

1.1 Objective

Driven by a commitment to enhance chest X-ray (CXR) interpretation, our project tackles inconsistencies and missed findings in diagnostics, particularly the 19% oversight in early lung cancer nodules. The goal is to develop an intelligent system for swift and accurate diagnosis, prioritizing patient care over time-consuming analyses. Building on the significance of CXRs in modern diagnostics and leveraging PACS-stored data, our initiative bridges the gap between radiological knowledge databases and deep learning frameworks. This synergy aims to propel CAD systems to new heights, promising groundbreaking improvements in diagnostic precision, especially crucial during global health crises like the COVID-19 pandemic.

1.2 Motivation

In India, the scarcity of radiologists burdens doctors with laborious chest X-ray reviews, increasing the risk of errors. AI-driven transformation in CXR interpretation holds the promise of elevating diagnostic accuracy, mitigating oversights, and enhancing patient outcomes. This is vital for addressing healthcare quality concerns and reducing malpractice risks. Chest X-rays play a crucial role in swiftly detecting respiratory conditions, and our project optimizes PACS-stored data for deep learning, aiming to revolutionize CAD systems. Amidst a global radiologist shortage, particularly evident during crises like COVID-19, our initiative focuses on improving disease localization accuracy, bridging expertise gaps, and delivering a powerful diagnostic tool for public health emergencies. [3]

1.3 Existing literature

The existing literature on chest X-ray analysis employing generated bounding boxes encompasses a diverse array of contributions. "Automatic lung segmentation in chest X-ray images using improved U-Net" by Wufeng Liu et al. (2022) [4] focuses on enhancing lung segmentation accuracy through a refined U-Net model. Ramkumar et al.'s (2020) [5] "Multi res U-Net based image segmentation of pulmonary tuberculosis using CT images" employs a multi-resolution U-Net for precise tuberculosis segmentation in CT scans. Kaviani et al.'s (2020) [3] study explores the application of AI to reduce missed findings in chest radiographs internationally.

Xiaosong Wang et al. (2017) [2] present the "ChestX-ray8" database, offering benchmarks for weakly-supervised classification and localization of thoracic diseases on a hospital scale. The foundational work by Huang et al. (2017) [6] on "Densely Connected Convolutional Networks" provides crucial insights into network architectures. Tamura and Yoshinaga's (2023) [7] "Segmentation-Based Bounding Box Generation for Omnidirectional Pedestrian Detection" contributes techniques for bounding box generation relevant to segmentation tasks.

These papers collectively contribute to the advancements in chest X-ray analysis, encompassing lung segmentation, tuberculosis detection, missed findings reduction, database creation, convolutional network architectures, and segmentation-based bounding box generation.

1.4 Expected Outcome

The expected outcome of our approach to chest X-ray analysis focusses on transformative advancement in medical diagnostics. By implementing advanced attention mechanisms from transformer models, we aim to significantly enhance disease localization accuracy, surpassing current ResNet-based models. Additionally, the incorporation of transfer learning with InceptionNet promises improved feature extraction, potentially elevating the overall classification performance in chest X-ray analysis. Rigorous refinement of the existing model further enhances disease detection accuracy and localization precision. The innovative bounding box generation, leveraging heatmaps and gradient maps, is expected to revolutionize disease-specific feature identification. This holistic strategy addresses inconsistencies in CXR interpretation, promising a paradigm shift toward intelligent, efficient, and precise healthcare diagnostics.

Chapter 2

Dataset and Data Prepocessing

2.1 Dataset (Initial findings)

The dataset employed for early findings in image segmentation comprises 800 samples, each with an image dimension of 1024x1024 pixels. Among these, 700 samples are accompanied by corresponding masks, enriching the dataset for supervised segmentation tasks. The data is strategically split for model training, validation, and testing purposes. Specifically, 630 samples are allocated for training, 70 for validation, and 100 for testing. This comprehensive dataset, with a balanced distribution across training and evaluation sets, serves as a robust foundation for exploring and assessing the efficacy of image segmentation models in early stages of development.

2.2 Dataset Overview

The dataset discussed consists of 108,948 medical imaging samples from the NIH CXR-8 dataset, each annotated with specific disease categories. Predominantly, 84,312 instances belong to the 'Normal' category, while abnormal cases exhibit varying frequencies.

Among the pathology classes, 'Infiltration' exhibits the highest frequency, with 10,317 samples, signifying its prominence in the dataset. Notably, 'Effusion', 'Mass' and 'Atelectasis' follow closely with 6,331, 6,046 and 5789 samples, respectively. The

classes 'Nodule', 'Pneumothorax', and 'Pneumonia' contribute moderately, showcasing a balanced representation of diverse abnormalities.

Furthermore, the dataset includes 1,000 samples with annotated bounding boxes, indicating a subset with localized abnormalities for better analysis. The images are standardized to a dimension of 1024x1024 pixels, ensuring consistency across the dataset.

In conclusion, this dataset offers a rich and diverse collection of medical images, with a focus on pulmonary pathologies. The distribution among classes provides valuable insights into the prevalence of specific conditions, enabling targeted research and analysis within the medical imaging domain.

2.3 Data Preprocessing

The preprocessing phase primarily consists of :

- Resizing images
- ToTensor Transformations
- Normalization

In the data preprocessing pipeline, images undergo a series of crucial steps to ensure uniformity and optimize compatibility with neural network models. Initially, all images are resized to a standardized format of 512x512 pixels, streamlining computational processes and establishing consistent input dimensions. Following this, the images are transformed into PyTorch tensors through the ToTensor operation, an essential conversion for seamless integration with neural networks. Subsequently, a normalization step is applied, involving the subtraction of mean values and division by standard deviation across the entire dataset. This normalization standardizes pixel values, effectively mitigating variations in brightness and contrast, which enhances model training stability. Together, these preprocessing steps create a cohesive and standardized dataset, providing a solid foundation for machine learning models by

facilitating efficient learning, generalization, and improved performance in handling diverse input patterns.

2.4 Data Splitting and Validation

The dataset is strategically split for classification, with 70% (76,264 samples) designated for training, 10% (10,894 samples) for validation, and 20% (21,790 samples) for testing. This distribution facilitates effective model training, fine-tuning through validation, and unbiased evaluation on an independent test set. Additionally, 1,000 annotated bounding boxes in the validation set specifically assess the efficacy of bounding box algorithms, enhancing the overall utility of the dataset for comprehensive model assessment and refinement.

Chapter 3

System Model

The methodology adopted involves comprehensive data preprocessing to enhance quality, followed by the implementation of a sophisticated model architecture for disease classification. Rigorous training refines the model's predictive capabilities. Post-classification, heatmap and bounding box generation further elucidate critical regions, providing a holistic approach to understanding and addressing medical imaging challenges.

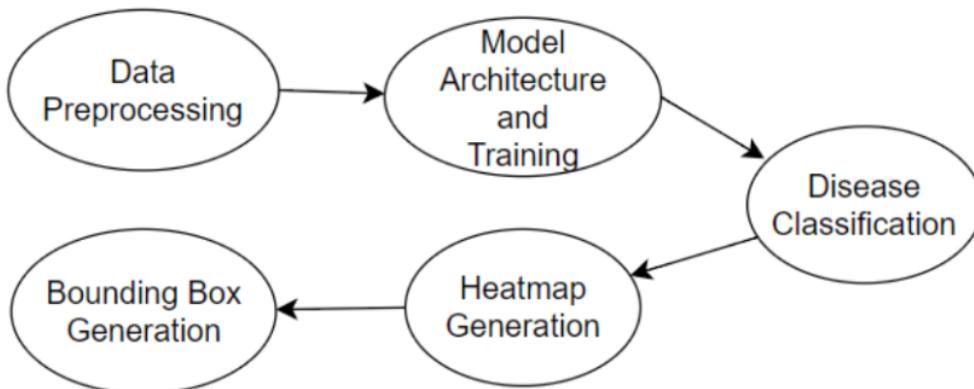


Fig. 3.1: *Methodology adopted*

3.1 U-Net

Significance of U-Net: The U-Net architecture plays a pivotal role in medical image analysis, particularly in lung segmentation, a critical step for accurate diagnostics. Its unique ability to capture both global context and intricate local details makes it highly effective. The model's precision in outlining lung boundaries within chest imaging sets the stage for subsequent diagnostic analyses. U-Net's significance lies in its contribution to enhancing the accuracy of diagnostic procedures, offering healthcare professionals a valuable tool for thorough and reliable assessments.

Architecture Overview: The U-Net architecture's design is characterized by a contracting encoder, bottleneck, and expansive decoder path. This structure allows it to capture hierarchical features while reconstructing segmented lung regions. Skip connections within the architecture facilitate the preservation of spatial details, contributing to the model's robustness. The contracting encoder employs convolutional layers with ReLU activations and max-pooling, crucial for extracting contextual information. The bottleneck layer refines abstract representations, maintaining critical information. The expansive decoder, utilizing up-convolutional layers and skip connections, gradually restores spatial resolution, aiding in the recovery of fine-grained details essential for accurate lung segmentation. The final layer employs a 1x1 convolution, generating pixel-wise predictions for segmented lung regions. This innovative architecture enables precise and efficient lung segmentation, showcasing its potential for broader applications in medical image analysis.

Conclusion and Implications: In conclusion, the U-Net architecture emerges as a robust and efficient solution for lung segmentation in medical images. Its successful application in delineating lung regions underscores its potential as a valuable asset ²² in the field of medical image analysis. The model's contribution to enhancing the precision and efficiency of diagnostic workflows suggests broader implications for its integration into healthcare practices, where accurate segmentation is crucial for informed decision-making and improved patient outcomes.

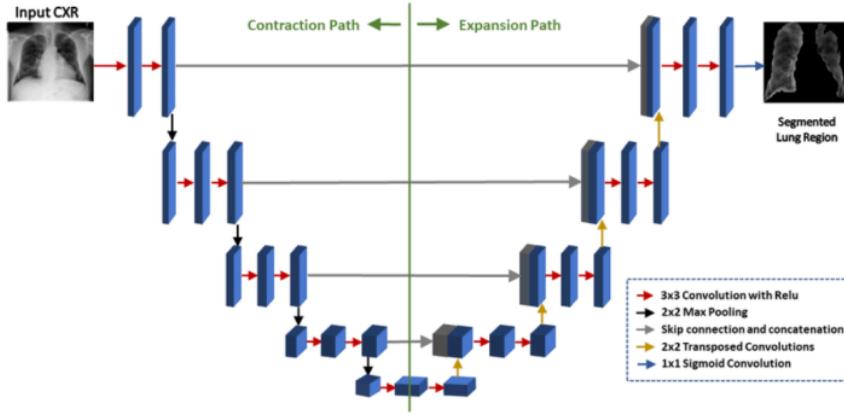


Fig. 3.2: Architecture of U-Net
[4]

3.2 DeepLabV3+

Significance of DeepLab-v3+ : DeepLab-v3+ stands out as a significant advancement in deep learning architecture, specifically tailored for semantic segmentation tasks like lung segmentation in medical images. Its prowess lies in accurately delineating intricate lung boundaries, providing healthcare professionals with a valuable tool for precise diagnostics. The model's ability to capture both local and global contextual information, particularly at various scales, makes it adept at handling complex lung structures and subtle boundaries, contributing significantly to its effectiveness in medical image analysis.

Architecture Overview: DeepLab-v3+ builds upon its predecessor by introducing an improved decoder module, enhancing segmentation performance. The architecture employs atrous convolutions, dilated convolutions, and a feature pyramid network (FPN) within a powerful encoder-decoder framework. Atrous convolutions, a key component, enable the model to enlarge its receptive field without inflating parameters, particularly useful in lung segmentation for capturing spatial dependencies across different scales. The ASPP module enhances multi-scale information capture, while the decoder, featuring depth-wise separable convolutions and FPN, refines segmentation results and ensures comprehensive contextual understanding. The final layer produces

pixel-wise predictions, outlining segmented lung regions.

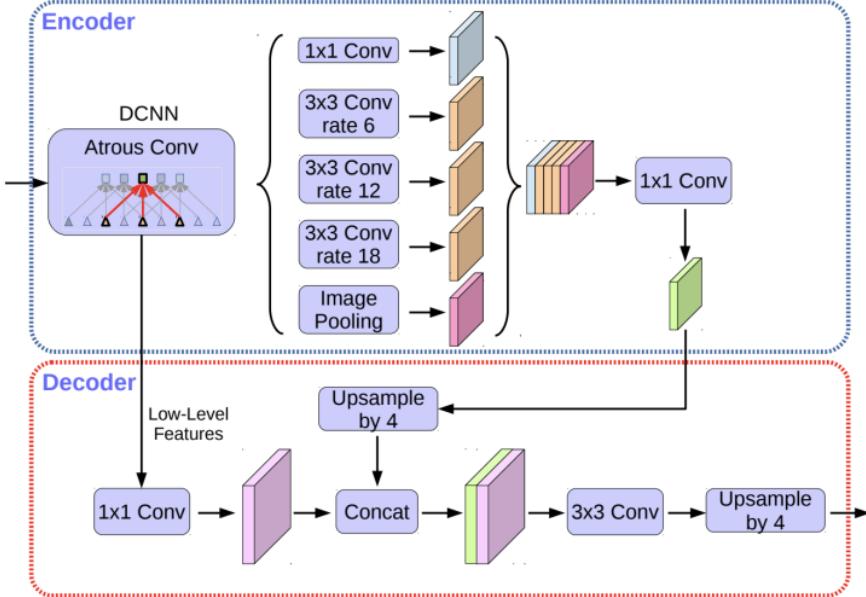


Fig. 3.3: Architecture of DeepLabv3+
[8]

Conclusion and Implications: In conclusion, DeepLab-v3+ emerges as a robust solution for lung segmentation in medical images. Its advanced features and innovative design contribute to accurate delineation of lung boundaries, enhancing diagnostic workflows in healthcare. The model's capability to handle complex structures and nuances in lung images suggests broad implications for its integration into medical practices, where precise segmentation is essential for informed decision-making and improved patient care. DeepLab-v3+ holds promise as a valuable asset in advancing medical image analysis applications.

3.3 Res U-Net

Significance of ResUNet in Medical Imaging: The Multi-resolution U-Net holds profound significance in the realm of medical image analysis, particularly in the context

of identifying pulmonary tuberculosis in CT images. Its unique capability to capture multi-scale information plays a pivotal role in delineating intricate structures within the lungs. This is crucial for detecting subtle abnormalities associated with tuberculosis, contributing significantly to improved diagnostic accuracy. The architecture's adaptability to binary segmentation specifically enhances its relevance in tuberculosis detection, making it a valuable asset in the field of medical imaging. 25

Architecture Overview: Inspired by the U-Net design, the Multi-resolution U-Net introduces enhancements in segmentation capabilities by incorporating multi-scale features. The architecture comprises an encoder-decoder structure with skip connections, preserving spatial details for accurate segmentation of lung regions. The encoder extracts hierarchical features, vital for capturing diverse textures and patterns indicative of pathological conditions, especially in the context of pulmonary tuberculosis detection. Notably, the Multi-resolution U-Net incorporates parallel encoding pathways, each focusing on a different level of detail, ensuring a comprehensive understanding of lung structures. The decoder, with skip connections, reconstructs segmented lung regions by recovering fine-grained details essential for precise segmentation.

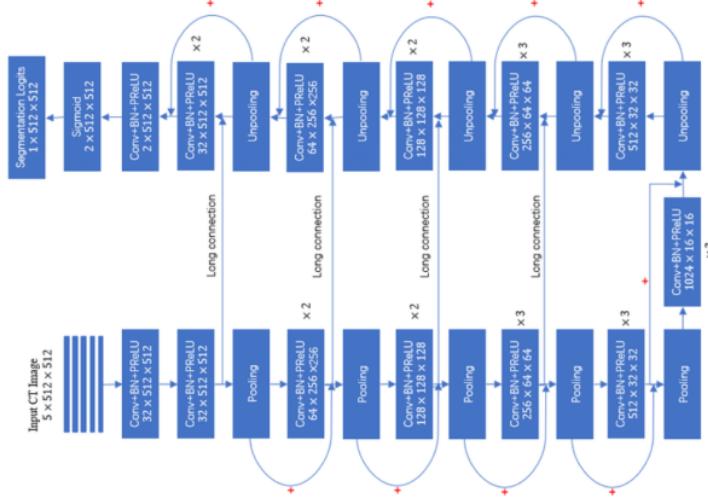


Fig. 3.4: Architecture of Res U-Net
[9]

²⁴ **Conclusion:** In conclusion, the Multi-resolution U-Net architecture emerges as a powerful and adaptive tool for lung segmentation, particularly in the nuanced domain of pulmonary tuberculosis detection using CT images. Its unique approach to multi-resolution feature extraction, coupled with skip connections, proves effective in precisely delineating lung structures. The architecture's contributions significantly enhance the diagnostic potential for tuberculosis in medical imaging, marking it as a valuable advancement in the ongoing efforts to improve disease detection and diagnosis accuracy.

3.4 GCN (Global Convolution Network)

Proposed Method-1: The model presented is an all-encompassing architecture tailored for lung segmentation and subsequent classification in chest X-ray images. By amalgamating a Global Context Module (GCM), Boundary Refinement Module (BRM), and a ResNet-152 backbone, the model strives to precisely identify key regions within the lungs and subsequently classify them with accuracy.[5]

Significance of GCN

The model's integration of GCM, BRM, and ResNet-152 forms a comprehensive solution for lung segmentation and classification. GCM enhances global contextual awareness, crucial for intricate pattern identification, while BRM refines boundaries for precise segmentation. The ResNet backbone ensures robust feature extraction, enhancing the model's effectiveness in image classification for pulmonary conditions.

Architecture: The model employs a ResNet-152 backbone for feature extraction, integrates a Global Context Module and Boundary Refinement Module, and utilizes ²⁰ upsampling for accurate lung segmentation in chest X-ray images.

Global Context Module (GCM): The GCM captures global contextual information across input feature maps, crucial for understanding complex relationships in spatially distant pixels. With instances applied at different scales, the GCM enriches the model's awareness of global patterns and structures, enhancing its segmentation capabilities.

Boundary Refinement Module (BRM): The BRM focuses on refining boundaries

within the input feature maps, contributing to more precise segmentation results. By utilizing consecutive convolutions with ReLU activation and a skip connection, the BRM enhances nuanced information at object or region boundaries, mitigating the vanishing gradient problem.

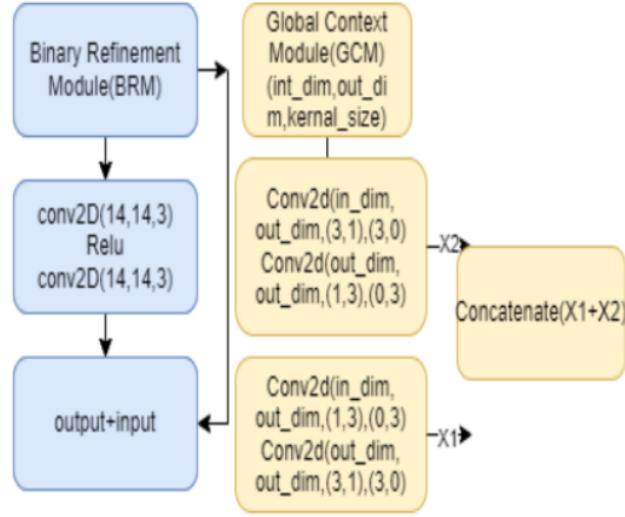


Fig. 3.5: Architecture of GCM and BRM

The model utilizes bilinear upsampling for enhanced spatial resolution, fusing upsampled feature maps with lower-layer features for accurate segmentation. The classifier, comprising fully connected layers, transforms features into classification outputs, yielding a segmentation map for identifying chest X-ray regions of interest.

Conclusion: In conclusion, the proposed model stands as a sophisticated and effective solution for lung segmentation and classification. Through the incorporation of advanced modules alongside a potent ResNet backbone, the model demonstrates significant potential in advancing diagnostic capabilities for pulmonary conditions within medical imaging applications.

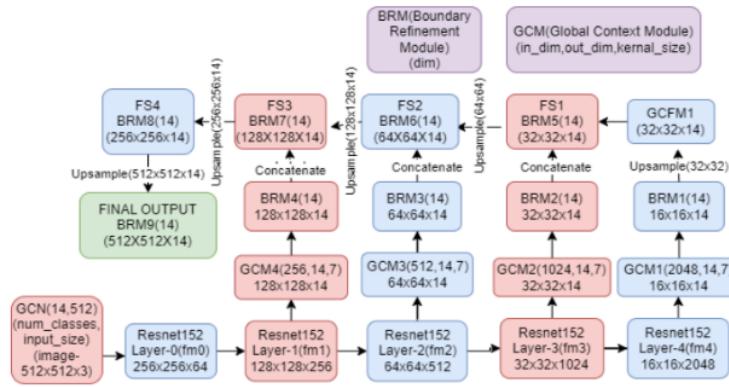


Fig. 3.6: Architecture of Segmentation Model

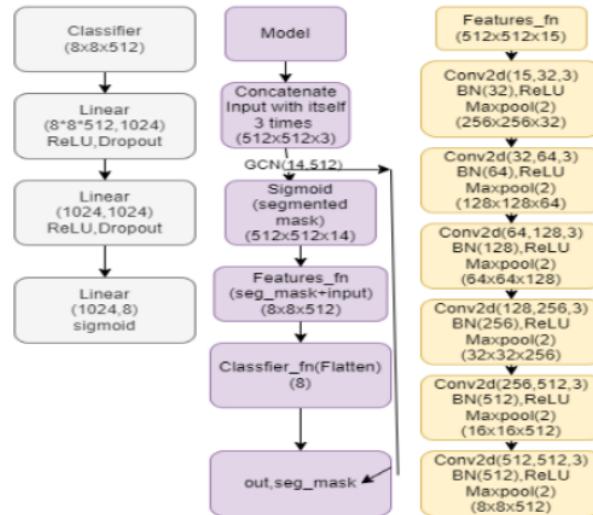


Fig. 3.7: Architecture of Global Convolution Network

3.5 DenseNet-121

Proposed Method-2: DenseNet121, is a renowned deep learning architecture for image classification, notably in medical imaging like chest X-ray disease classification. Its dense connections and deep structure contribute to its popularity and effectiveness in this domain.[6]

Significance if DenseNet 121

DenseNet121 holds significance in medical imaging for disease classification due to its unique features. Its dense connectivity allows efficient feature reuse, crucial for fine-grained pattern recognition. Smooth gradient flow overcomes vanishing gradient issues, and parameter efficiency reduces redundancy, leading to accurate disease classification. The model's pre-trained versions enable effective transfer learning, especially beneficial in scenarios with limited labeled medical datasets.

DenseNet121 Architecture DenseNet's architecture introduces innovative elements such as dense blocks, where each layer receives input from all preceding layers, promoting extensive feature reuse and enhancing information flow. Transition layers, inserted between dense blocks, reduce spatial dimensions and control parameters. Bottleneck layers within dense blocks employ 1x1 and 3x3 convolutions, optimizing computational efficiency. Global average pooling replaces fully connected layers, reducing parameters and preventing overfitting. While dropout layers may be added for regularization, the original DenseNet121 design focuses on dense connectivity, efficient transitions, and global average pooling, collectively forming a powerful architecture known for feature reuse, streamlined information flow, and parameter efficiency in tasks like disease classification in medical images.

Conclusion: In conclusion, DenseNet121 excels in disease classification for chest X-rays, offering efficient feature reuse, gradient flow, and parameter efficiency. Its significance lies in high accuracy, transfer learning, and interpretability, making it popular in medical image analysis. Practical implementation requires addressing challenges like data augmentation and class imbalance for successful deployment in healthcare applications.

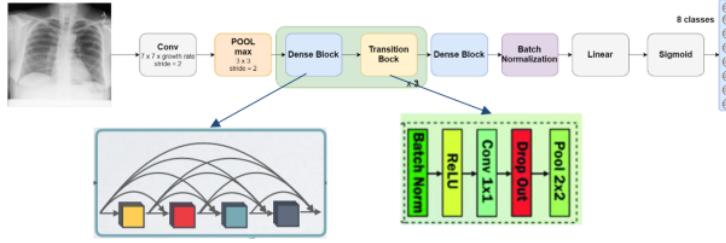


Fig. 3.8: Architecture of DenseNet Model

3.6 Bounding Box Generation

This study investigates the imperative need for precise disease localization in medical imaging for enhanced diagnostics and treatment planning. We delve into two sophisticated techniques, Class Activation Maps (CAM) and Grad-CAM, adopting methodologies from Zhou et al. and Grad-CAM to refine disease prediction and facilitate bounding box localization. Our approach offers a comprehensive solution, improving localization accuracy without necessitating extensive model modifications.

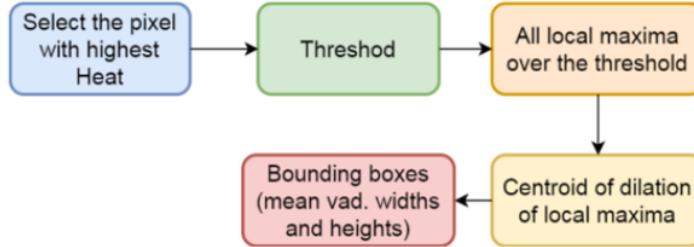
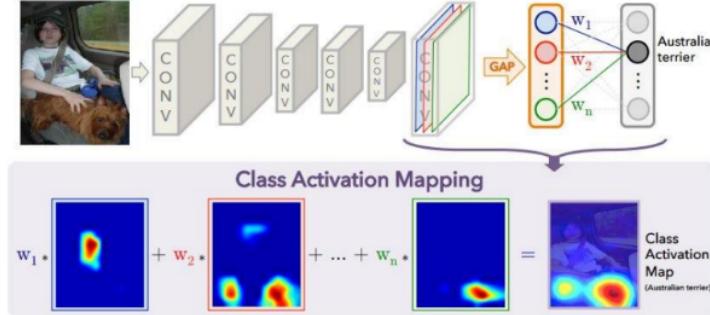


Fig. 3.9: Bounding Box work flow

CAM (Class Activation Map): CAM is implemented through a linear combination of feature maps preceding the Global Average Pooling (GAP) and dense layers in the model's final section. Adopting Zhou et al.'s methodology, it generates heat-maps for precise disease localization. Particularly effective in models with global average and dense layers, CAM facilitates discriminative localization without re-training. The resulting heat-maps enhance interpretability, aiding clinicians in identifying critical regions and comprehending the model's decision-making process, thereby contributing

to improved diagnostic understanding and decision support in medical settings.



¹⁵
Fig. 3.10: Class Activation Mapping

Grad-CAM (Gradient-weighted Class Activation Mapping): In contrast to CAM, Grad-CAM utilizes gradients flowing into the final convolutional layer for disease localization. ²³ This method generates a coarse map, highlighting key image regions for disease prediction. Remarkably, Grad-CAM is applicable to various CNN-based architectures without requiring modifications to the premodel structure. Its flexibility streamlines implementation, minimizing the need for extensive model adjustments. The absence of re-training enhances Grad-CAM's efficiency and generalizability in disease localization tasks.

Bounding Box Localization: Post heat-map generation, a precise bounding box localization approach is applied. Utilizing a threshold based on the global peak value, local maximum points are selected, refined with maximum and minimum filters. Dilation accumulates candidate points, and their centroid becomes the bounding box center. A fixed-size strategy is adopted, adjusting thresholds for disease classes. Mean width and height from constructed bounding boxes ensure consistent and accurate disease region representation, contributing to enhanced localization precision.

Conclusion: CAM and Grad-CAM, combined with an inventive bounding box localization strategy, establish a robust framework for disease prediction and localization in medical imaging. Their seamless integration enhances the interpretability of deep learning models, fostering a transparent and clinically relevant decision-making process. The proposed methodology holds promise for advancing medical image analysis,

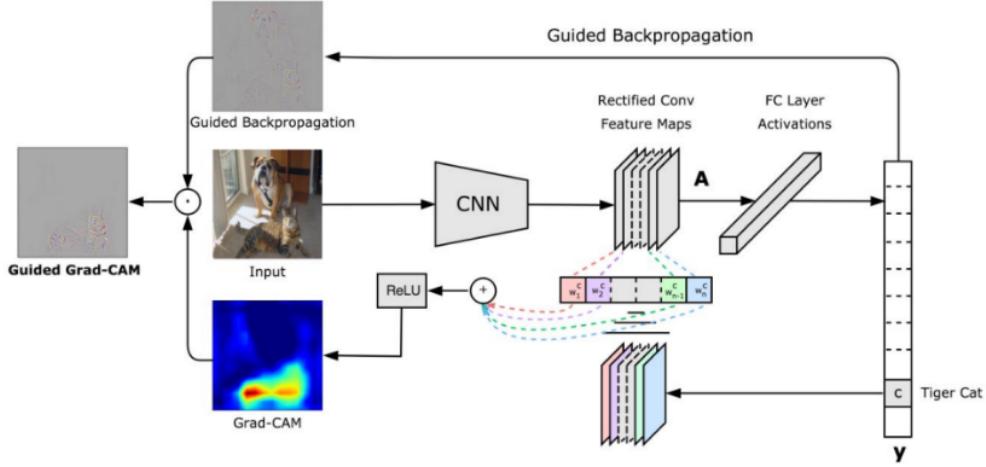


Fig. 3.11: *Gradient-weighted Class Activation Mapping*

contributing to enhanced diagnostic accuracy and patient care.

3.7 Weighted Binary Cross-Entropy

In binary classification, Weighted Binary Cross-Entropy is introduced to counter class imbalances by assigning distinct weights to positive and negative samples. This tailored approach ensures more balanced learning, especially when one class dominates, prioritizing accurate classification of the minority class for meaningful model training.

Mathematically, Weighted Binary Cross-Entropy is calculated using the formula:

$$\text{Weighted BCE Loss} = -\frac{1}{N} \sum_{i=1}^N (w_{\text{pos}} \cdot y_i \cdot \log(p_i) + w_{\text{neg}} \cdot (1 - y_i) \cdot \log(1 - p_i)) \quad (3.1)$$

Where :

N is the total number of samples.

y_i is the true label of the i -th sample (0 or 1).

p_i is the predicted probability of the i -th sample belonging to the positive class.

w_{pos} and w_{neg} are the weights assigned to the positive and negative classes

Chapter 4

Results

Our AUC scores are somewhat similar to or better than existing benchmarks, highlighting the effectiveness of our approach in medical image analysis

- Comparison of loss values and dice coefficient values has been shown for UNet model in fig. 4.1. From the plot we can see that training dataset is showing better results compared to training dataset.

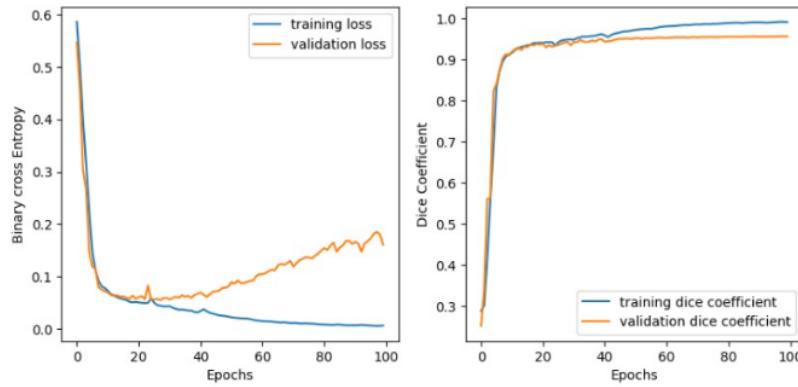


Fig. 4.1: Loss and dice coefficient values for UNet model

- Comparison of loss values and dice coefficient values has been shown for ResUNet model in fig. 4.2. From the plot we can see that training dataset is showing better results compared to training dataset.

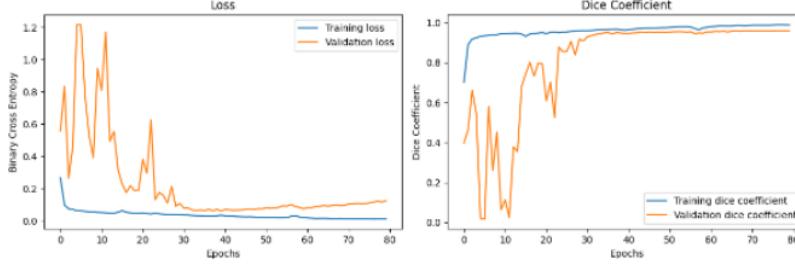


Fig. 4.2: Loss and dice coefficient values for ResUNet model

- Our AUC scores for image segmentation are shown in table 4.1 , The UNet model exhibits Dice Score of 0.94, showcasing its robust segmentation capabilities. Deeplabv3+ demonstrates accuracy of 0.88, while ResUnet surpasses all with an impressive Dice Score of 0.96, indicating superior segmentation precision. These results highlight the effectiveness of ResUnet in accurately delineating objects in images for advanced applications.

Table 4.1: Result Comparisons for image segmentation

Model Name	Dice Score
UNet	0.94
Deeplab	0.88
ResUnet	0.96

- Loss and AUC Values for Global Convolution Model is shown in fig. 4.3.

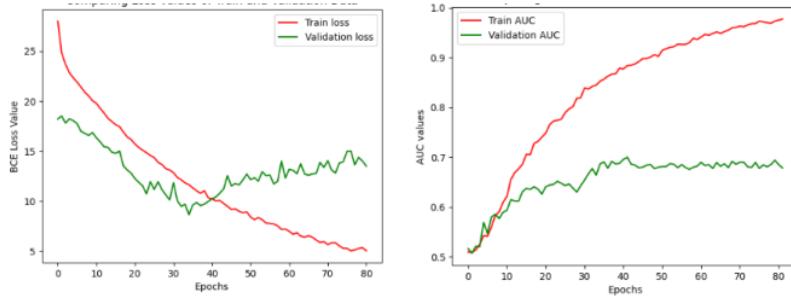


Fig. 4.3: Loss and AUC values for GCN model

- From the results shown in figure 4.5 we can see that Model-1 demonstrates competitive performance across diseases, with notable AUC scores. While not consistently outperforming Model-2, it presents a strong alternative with respectable accuracies, making it a reliable choice for pathology detection tasks.

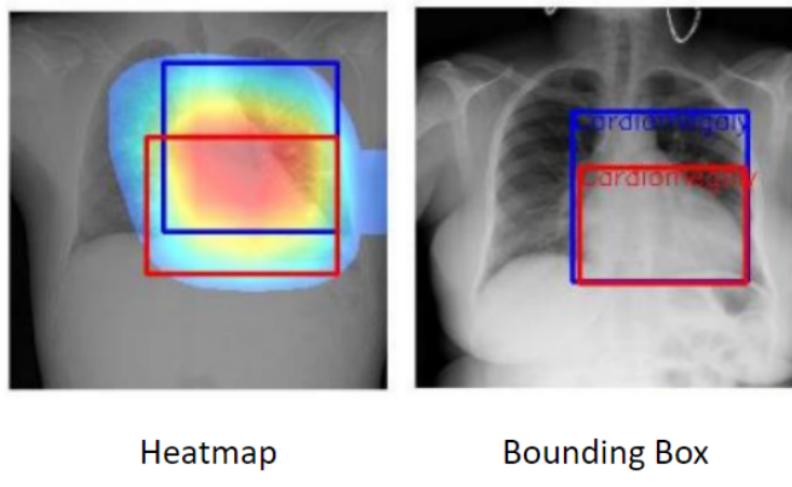
Model-2 consistently excels across diseases, exhibiting the highest AUC scores. Its robust performance suggests superior pathology detection capabilities, making it a compelling choice for medical image analysis. Its effectiveness surpasses existing models in the dataset.

In comparison, Model-2 consistently outperforms Model-1 and other existing models in the dataset, showcasing its effectiveness in pathology detection. Its higher AUC scores across various diseases highlight its potential for enhanced diagnostic accuracy and reliability in medical imaging applications.

Pathology	Wang et at. (2017)	Yao et al. (2017)	Pranav Rajpurkar (2017)	Model-1	Model-2 (DenseNet-121)
Atelectasis	0.716	0.772	0.8094	0.7249	0.8486
Cardiomegaly	0.807	0.904	0.9248	0.8373	0.9479
Effusion	0.784	0.859	0.8638	0.8441	0.8884
Infiltration	0.609	0.695	0.7345	0.5956	0.6630
Mass	0.706	0.792	0.8676	0.5270	0.8147
Nodule	0.671	0.717	0.7802	0.7784	0.8066
Pneumonia	0.633	0.713	0.7680	0.6718	0.7228
Pneumothorax	0.806	0.841	0.8887	0.6063	0.8996

Fig. 4.4: AUC score comparison for different diseases with existing works[1] [2]

- The Intersection over Union (IoU) for bounding boxes is 0.2530, reflecting the overlap between predicted and ground truth boxes. Figure 4.6 displays the generated bounding box output, illustrating the model's precise localization. This emphasizes the accuracy of our approach in delineating regions of interest in the study.



Heatmap Bounding Box

Fig. 4.5: Showing the final output with bounding box

Chapter 5

Future Work

In future work, the focus lies on pushing the boundaries of chest x-ray analysis. Innovations include exploring novel architectures, optimizing model interpretability, and adapting to evolving medical imaging technologies for enhanced diagnostic capabilities. The following avenues represent key focus areas for future work and improvements:

Advanced Attention Mechanisms Integration: Integrate advanced attention mechanisms from transformer models to significantly improve disease localization accuracy in chest x-ray analysis. This enhancement aims to surpass the current capabilities of models relying on resnet-based architectures, offering superior performance in pinpointing disease regions within the images.

Transfer Learning with InceptionNet: Utilize transfer learning techniques with InceptionNet to boost feature extraction capabilities. This approach seeks to elevate the overall performance of chest x-ray analysis by harnessing the robust feature representation offered by InceptionNet, potentially enhancing the model's ability to discern complex patterns.

Rigorous Refinement for Enhanced Disease Detection: Execute a thorough refinement process on the existing model, prioritizing improvements in both classification and localization aspects. The goal is to achieve heightened accuracy in disease detection and precise localization within chest x-ray images, ensuring a more robust and reliable diagnostic capability.

Chapter 6

Timeline

From August to September 2023, we conducted a literature review and finalized the problem statement. In October 2023, models for lung segmentation were trained. Subsequently, from November to December 2023, models for chest x-ray classification and bounding box generation were created. In January–February 2024, transformer and inception models were integrated, and March–April 2024 focused on enhancing bounding box generation.



Fig. 6.1: Detailed timeline

Bibliography

- [1] ³ Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, Andrew Y. Ng, “CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning,” 2017.
- [2] ⁶ Xiaosong Wang, Yifan Peng , Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, Ronald M. Summers 1 , “ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases,” 2017.
- [3] ¹ Parisa Kaviani, Mannudeep K Kalra, Subba R Digumarthy, Reya V Gupta, Giridhar Dasegowda , Ammar Jagirdar, Salil Gupta, Preetham Putha, Vidur Mahajan, Bhargava Reddy, Vasanth K Venugopal, Manoj Tadepalli, Bernardo C Bizzo, Keith J Dreyer , ‘Frequency of Missed Findings on Chest Radiographs (CXRs) in an International, Multicenter Study: Application of AI to Reduce Missed Findings,’ 2020.
- [4] ⁸ Y. Y. W. W. J. D. . L. Y. Wufeng Liu, Jiaxin Luo, *Automatic lung segmentation in chest X-ray images using improved U-Net*. Nature, 2022.
- [5] ⁹ MO Ramkumar, D Jayakumar, R Yogesh, “Multi res U-Net based image segmentation of pulmonary tuberculosis using CT images,” 2020.
- [6] ¹¹ Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger , “Densely Connected Convolutional Networks,” 2017.
- [7] ⁵ Masato Tamura, Tomoaki Yoshinaga , “Segmentation-Based Bounding Box Generation for Omnidirectional Pedestrian Detection,” 2023.
- [8] ⁷ Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, Hartwig Adam, “Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation,” 2018.
- [9] ⁵ Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, Antonio Torralba, “Learning Deep Features for Discriminative Localization,” 2015.

Report

ORIGINALITY REPORT

10%
SIMILARITY INDEX

9%
INTERNET SOURCES

8%
PUBLICATIONS

7%
STUDENT PAPERS

PRIMARY SOURCES

- | | | |
|----------|---|-----------|
| 1 | www.mdpi.com
Internet Source | 1% |
| 2 | Submitted to Indian Institute of Technology
Student Paper | 1% |
| 3 | arxiv.org
Internet Source | 1% |
| 4 | Submitted to Georgia State University
Student Paper | 1% |
| 5 | export.arxiv.org
Internet Source | 1% |
| 6 | web.archive.org
Internet Source | 1% |
| 7 | Qixian Zhou, Xiaodan Liang, Ke Gong, Liang Lin. "Adaptive Temporal Encoding Network for Video Instance-level Human Parsing", 2018 ACM Multimedia Conference on Multimedia Conference - MM '18, 2018
Publication | 1% |
| 8 | assets.researchsquare.com
Internet Source | |

<1 %

-
- 9 Paola Barra, Alessia Auriemma Citarella, Giosu  Orefice, Modesto Castrill n-Santana, Angelo Ciaramella. "LOTS: Litter On The Sand dataset for litter segmentation", 2023 18th International Conference on Machine Vision and Applications (MVA), 2023 <1 %
Publication
-
- 10 pubmed.ncbi.nlm.nih.gov <1 %
Internet Source
-
- 11 zihua-liu.github.io <1 %
Internet Source
-
- 12 Shubhi Miradwal, Waquas Mohammad, Anvi Jain, Fawwaz Khilji. "Chapter 25 Lesion Segmentation in Skin Cancer Detection Using UNet Architecture", Springer Science and Business Media LLC, 2023 <1 %
Publication
-
- 13 digitalcommons.mtu.edu <1 %
Internet Source
-
- 14 Submitted to University of Strathclyde <1 %
Student Paper
-
- 15 Submitted to University of Surrey <1 %
Student Paper
-
- dblp.uni-trier.de

16	Internet Source	<1 %
17	www.chemometrie.com Internet Source	<1 %
18	dokumen.pub Internet Source	<1 %
19	elibrary.stipram.ac.id Internet Source	<1 %
20	spectrum.library.concordia.ca Internet Source	<1 %
21	vtechworks.lib.vt.edu Internet Source	<1 %
22	www.frontiersin.org Internet Source	<1 %
23	Behrouz Ebrahimi, David Le, Mansour Abtahi, ALBERT DADZIE et al. "Assessing Spectral Effectiveness in Color Fundus Photography for Deep Learning Classification of Retinopathy of Prematurity", Optica Publishing Group, 2023 Publication	<1 %
24	Fantong Kong, Francesco Picetti, Vincenzo Lipari, Paolo Bestagini, Stefano Tubaro. "Deep prior-based seismic data interpolation via multi-res U-net", SEG Technical Program Expanded Abstracts 2020, 2020	<1 %

- 25 "Image Analysis for Moving Organ, Breast, and Thoracic Images", Springer Science and Business Media LLC, 2018 <1 %
- Publication
-
- 26 Jiaqi Shao, Shuwen Chen, Jin Zhou, Huisheng Zhu, Ziyi Wang, Mackenzie Brown. "Application of U-Net and Optimized Clustering in Medical Image Segmentation: A Review", Computer Modeling in Engineering & Sciences, 2023 <1 %
- Publication
-
- 27 Nahian Siddique, Sidike Paheding, Colin P. Elkin, Vijay Devabhaktuni. "U-net and its variants for medical image segmentation: A review of theory and applications", IEEE Access, 2021 <1 %
- Publication
-

Exclude quotes

Off

Exclude matches

Off

Exclude bibliography

Off