Revista Latinoamericana de la Papa
Vol.29, No.1, 2025

ISSN 1019-6609
eISSN 1853-4961

Anatomy-Aware Deep Learning Model (Anatomy-XNet) for Thoracic Disease Detection in Chest X-rays

Dr. Sanjay Kumar Sharma
Assistant Professor, Department of Anatomy,
Government Medical College Sriganganagar, Rajasthan.
Email: drsanjaybkn56@gmail.com

ABSTRACT

Deep learning has shown great promise in the automatic detection of thoracic diseases using chest X-rays (CXR). However, most existing models lack anatomical awareness, often treating the entire image uniformly without accounting for organ-specific localization, which can lead to false positives and reduced interpretability. This paper presents Anatomy-XNet, a novel anatomy-aware convolutional neural network (CNN) that integrates anatomical priors to enhance thoracic disease classification. The model incorporates an Anatomy-Aware Attention Module (A^3M) and Probabilistic Weighted Average Pooling (PWAP) to focus on key anatomical structures such as the lungs, heart, and diaphragm, using organ-level annotations. We evaluated Anatomy-XNet on three major public datasets: NIH ChestX-ray14, CheXpert, and MIMIC-CXR, achieving AUC scores of 85.78%, 92.07%, and 84.04%, respectively. Compared to baseline DenseNet-121 models, Anatomy-XNet consistently outperformed in both classification accuracy and localization precision. The model also demonstrated improved interpretability through heatmap visualization, aligning closely with expert annotations. These results underscore the importance of incorporating anatomical context in medical image analysis and pave the way for more accurate and clinically useful AI tools in radiology.

KEYWORDS

Deep Learning, Chest X-ray, Thoracic Disease, Anatomy-Aware Model, Convolutional Neural Network (CNN), Attention Mechanism, Medical Imaging, Disease Classification, Artificial Intelligence in Radiology, Interpretability in AI.

1. Introduction

Thoracic diseases, including pneumonia, pneumothorax, pleural effusion, cardiomegaly, and pulmonary edema, are among the leading causes of morbidity and mortality worldwide. Chest X-rays (CXR) are one of the most commonly used imaging modalities for the initial assessment and diagnosis of these conditions due to their low cost, accessibility, and rapid acquisition. However, interpreting CXRs is a complex task, often requiring expert radiological knowledge to distinguish subtle abnormalities that may be indicative of disease. The global shortage of trained radiologists, particularly in low-resource settings, has fueled the development of artificial intelligence (AI) systems capable of automating the detection and classification of thoracic diseases in CXRs.

In recent years, deep learning—particularly convolutional neural networks (CNNs)—has demonstrated impressive performance in medical image analysis. Models such as DenseNet, ResNet, and EfficientNet have been successfully trained on large-scale datasets like NIH ChestX-ray14 and

ISSN 1019-6609
eISSN 1853-4961

CheXpert, achieving near-radiologist level accuracy in some tasks. However, despite these advances, most existing models treat the chest X-ray as a homogenous image without incorporating knowledge of human anatomy. This lack of anatomical context limits the model's ability to focus on disease-relevant regions, often resulting in poor localization, false positives, and reduced interpretability. For example, a model might detect a suspected consolidation in non-pulmonary areas, which is clinically implausible.

Radiologists, by contrast, interpret CXRs by systematically examining specific anatomical structures—the lungs, heart, diaphragm, mediastinum, and pleura—and relating abnormalities to these regions. Incorporating this anatomical reasoning into deep learning models could significantly enhance both the accuracy and clinical utility of AI-driven diagnostics. This forms the central motivation for our study: to develop a deep learning model that not only performs classification tasks but also embeds anatomical awareness in its feature extraction and decision-making processes.

We propose **Anatomy-XNet**, an anatomy-aware deep learning model specifically designed for thoracic disease detection in CXRs. The model introduces two key innovations: the **Anatomy-Aware Attention Module (A³M)** and **Probabilistic Weighted Average Pooling (PWAP)**. A³M directs the model's attention to anatomically relevant regions, guided by organ-level annotations, while PWAP ensures that features extracted from critical areas are given greater importance during the classification phase. These modules work in tandem to align the model's focus with clinical reasoning, improving both diagnostic accuracy and interpretability.

Anatomy-XNet is trained and evaluated on three widely used benchmark datasets: NIH ChestX-ray14, CheXpert, and MIMIC-CXR. We assess the model's performance in terms of area under the ROC curve (AUC), precision-recall metrics, and localization accuracy via saliency map analysis. Additionally, we compare its performance with standard CNN architectures lacking anatomical awareness to demonstrate the added value of integrating anatomical priors.

The findings from this research not only offer a more clinically aligned AI tool for CXR interpretation but also highlight the broader importance of incorporating domain-specific knowledge into medical AI models. By aligning deep learning with human anatomical reasoning, Anatomy-XNet represents a step forward in building interpretable, accurate, and trustworthy AI systems for radiology.

2. Materials and Methods

1. Overview

This study presents **Anatomy-XNet**, a deep learning framework designed to improve thoracic disease detection in chest X-rays (CXRs) by incorporating anatomical awareness. The architecture enhances diagnostic precision and interpretability by embedding organ-level focus mechanisms into standard convolutional neural networks (CNNs). The following subsections outline the datasets used, preprocessing steps, model architecture, training procedure, and evaluation metrics.

2. Datasets

Anatomy-XNet was trained and evaluated on three publicly available chest X-ray datasets:

- **NIH ChestX-ray14**: Comprises 112,120 frontal-view X-ray images of 30,805 unique patients labeled with 14 thoracic disease categories.
- **CheXpert**: Contains 224,316 chest radiographs of 65,240 patients labeled for 14 observations including uncertain labels.

ISSN 1019-6609
eISSN 1853-4961

- **MIMIC-CXR**: Consists of over 370,000 CXR studies from Beth Israel Deaconess Medical Center, linked with free-text radiology reports.

To embed anatomical context, we used a subset of each dataset with available **organ-level annotations** (e.g., lung fields, cardiac silhouette, diaphragm), either provided directly or inferred using an auxiliary segmentation model.

3. Preprocessing

All images were resized to **224×224 pixels** for uniformity and computational efficiency. Data augmentation techniques including rotation (±10°), horizontal flipping, and brightness normalization were applied to increase robustness. Pixel intensity was normalized to the [0,1] range. For organ-aware attention, binary masks representing anatomical regions were applied as spatial priors during training.

Uncertain labels (present in CheXpert) were handled using **U-Zero policy**, where uncertain findings were treated as negative to minimize label noise.

4. Model Architecture

Anatomy-XNet builds upon the **DenseNet-121** backbone due to its strong baseline performance in medical imaging tasks. Two key modules were added:

- **Anatomy-Aware Attention Module (A^3M)**: Applies a learned attention map over predefined anatomical regions. This map highlights clinically relevant areas such as the lungs, heart, and diaphragm, ensuring the model allocates more weight to features originating from these structures.
- **Probabilistic Weighted Average Pooling (PWAP)**: Replaces global average pooling with a probabilistic mechanism that assigns higher weight to activations localized within anatomically significant regions, improving interpretability and localization accuracy.

The output layer includes **14 sigmoid-activated neurons** corresponding to the 14 disease labels. A multi-label binary cross-entropy loss function was used for training.

5. Training Procedure

Training was performed using **Adam optimizer** with an initial learning rate of 0.001, decayed using a cosine annealing scheduler. The model was trained for **30 epochs** with a batch size of **32** on an NVIDIA Tesla V100 GPU. Early stopping based on validation loss was applied to avoid overfitting.

6. Evaluation Metrics

Model performance was evaluated using:

- **Area Under the Receiver Operating Characteristic Curve (AUC)** for each disease class.
- **Localization Accuracy**, assessed using Grad-CAM heatmaps and intersection-over-union (IoU) with ground truth annotations.
- **Precision, Recall, and F1-score** for overall classification performance.
- **Statistical Significance** between Anatomy-XNet and baseline DenseNet was tested using paired t-tests.

ISSN 1019-6609
eISSN 1853-4961

3. Results

This section presents the performance outcomes of **Anatomy-XNet** compared to a baseline **DenseNet-121** model trained under identical conditions but without anatomical awareness. The models were evaluated on three benchmark datasets: **NIH ChestX-ray14**, **CheXpert**, and **MIMIC-CXR**. Key metrics include **Area Under the Curve (AUC)**, **Precision**, **Recall**, **F1-Score**, and **Localization Accuracy**, with particular attention to performance on disease detection and interpretability via anatomical attention.

1. Classification Performance

Anatomy-XNet outperformed the baseline DenseNet-121 in average AUC across all datasets. The largest performance gain was observed in diseases that are highly dependent on localized anatomical cues (e.g., cardiomegaly, effusion, consolidation).

Table 1: Comparison of Average AUC Scores Across Datasets

Dataset	DenseNet-121 (AUC)	Anatomy-XNet (AUC)	Δ AUC
NIH ChestX-ray14	82.1%	85.78%	+3.68%
CheXpert	89.0%	92.07%	+3.07%
MIMIC-CXR	80.2%	84.04%	+3.84%

Anatomy-XNet achieved **statistically significant improvements** ($p < 0.01$) across all datasets. The model consistently showed higher sensitivity (recall) without compromising specificity (precision).

2. Per-Class Performance

To better understand the model's strengths, we analyzed AUC values by disease class on the CheXpert dataset.

Table 2: Disease-wise AUC Comparison on CheXpert Dataset

Condition	DenseNet-121 AUC	Anatomy-XNet AUC
Cardiomegaly	87.6%	91.2%
Consolidation	83.1%	86.5%
Pleural Effusion	88.9%	92.3%
Pneumothorax	89.5%	91.6%
Edema	85.3%	88.7%
Atelectasis	84.2%	86.4%

These improvements reflect the benefit of anatomical attention, particularly in conditions with clear spatial dependencies.

3. Localization Accuracy

To assess interpretability, we used **Grad-CAM** to visualize heatmaps of model attention. Heatmaps from Anatomy-XNet were more anatomically focused, aligning well with expert annotations. We

ISSN 1019-6609
eISSN 1853-4961

quantified this using **Intersection-over-Union (IoU)** between Grad-CAM masks and organ-specific bounding boxes.

- **Average IoU (Lung-based pathologies):**
    - DenseNet-121: **0.41**
    - Anatomy-XNet: **0.58** ($p < 0.01$)

This 41% relative improvement in localization accuracy confirms the effectiveness of anatomy-aware attention in enhancing model focus and transparency.

4. Error Analysis

A review of misclassified samples revealed that Anatomy-XNet was less likely to activate on irrelevant anatomical regions (e.g., soft tissues outside the thoracic cavity), reducing false positives. For example, false consolidation signals over the abdomen were frequent in the baseline model but markedly reduced in Anatomy-XNet predictions.

5. Computational Performance

Despite the added modules, Anatomy-XNet maintained computational efficiency:

- **Training time per epoch**: 7.4 min (vs. 6.8 min for DenseNet)
- **Inference time per image**: 0.18 sec (vs. 0.15 sec)

This marginal increase is acceptable considering the interpretability and accuracy gains.

Summary

Anatomy-XNet demonstrates superior diagnostic accuracy and anatomical localization across three benchmark datasets. The incorporation of anatomical priors through attention mechanisms significantly enhances the model's clinical relevance and interpretability, making it a robust candidate for AI-assisted radiology workflows.

4. Discussion

The results demonstrate that incorporating anatomical awareness into deep learning models can significantly improve thoracic disease detection from chest X-rays. Anatomy-XNet consistently outperformed the baseline DenseNet-121 across multiple datasets, highlighting the importance of integrating organ-specific context in medical image analysis. By focusing the model's attention on clinically relevant anatomical regions such as the lungs, heart, and diaphragm, Anatomy-XNet was better able to localize disease manifestations, leading to improved classification accuracy and reduced false positives.

The improvements in localization accuracy, evidenced by higher Intersection-over-Union scores, also enhance model interpretability—a crucial factor for clinical adoption. Radiologists are more likely to trust AI predictions when attention maps align with known anatomical landmarks, making Anatomy-XNet's outputs more transparent and actionable.

Moreover, the model's robustness across three diverse datasets with varying label quality and patient populations demonstrates its generalizability. The slight computational overhead introduced by the attention modules is justified by the performance gains and interpretability benefits, ensuring the model remains practical for real-world deployment.

ISSN 1019-6609
eISSN 1853-4961

However, this study has limitations. The reliance on available organ-level annotations, either manually labeled or inferred, may restrict applicability where such data are unavailable. Additionally, while improvements were statistically significant, they are moderate, suggesting room for further optimization, such as incorporating temporal data from longitudinal studies or integrating clinical metadata.

Future work could explore extending anatomy-aware mechanisms to other imaging modalities and diseases, as well as investigating personalized models that adapt attention based on individual patient anatomy. Overall, Anatomy-XNet represents a promising step toward more accurate, interpretable, and clinically aligned AI in radiology.

5. Conclusion

This study presents **Anatomy-XNet**, an anatomy-aware deep learning model that enhances thoracic disease detection in chest X-rays by integrating anatomical priors into its attention mechanisms. Our results demonstrate that embedding organ-specific focus improves both classification accuracy and localization precision across multiple large-scale datasets, outperforming conventional CNN architectures. The improved interpretability, evidenced by more anatomically relevant attention maps, increases clinical trust and usability. While the model introduces a slight computational overhead, the benefits in diagnostic performance and transparency underscore its potential for real-world application in radiology workflows. Future research should aim to broaden anatomical context incorporation, explore other imaging modalities, and refine attention mechanisms to further optimize performance. Overall, Anatomy-XNet signifies a meaningful advancement toward AI systems that align more closely with clinical reasoning, offering promise for improved diagnostic support in thoracic imaging.

References

[1] 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.

[2] Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). ChestX-ray8: Hospitalscale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2097-2106.

[3] Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., ... & Ng, A. Y. (2019). CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. Proceedings of the AAAI Conference on Artificial Intelligence, 33(01), 590-597.

[4] Johnson, A. E. W., Pollard, T. J., Berkowitz, S. J., Greenbaum, N. R., Lungren, M. P., Deng, C. Y., ... & Mark, R. G. (2019). MIMIC-CXR, a large publicly available database of labeled chest radiographs. Scientific Data, 6(1), 317.

[5] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE International Conference on Computer Vision, 618-626.

[6] Rajpurkar, P., Irvin, J., Ball, R. L., Zhu, K., Yang, B., Mehta, H., ... & Ng, A. Y. (2018). Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Medicine, 15(11), e1002686.

[7] Guendel, S., Perez, A., Hoebel, K., Allen, B., Yao, J., & Rubin, D. L. (2018). Learning to recognize abnormalities in chest X-rays with location-aware dense networks. Medical Imaging with Deep Learning.

ISSN 1019-6609
eISSN 1853-4961

[8] Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology, 284(2), 574-582.

[9] Wang, H., Zhang, Y., Xu, Y., Liang, J., & Wang, Y. (2020). Anatomy-aware multi-label classification of chest X-ray images using region-based convolutional networks. IEEE Access, 8, 211429-211438.

[10] Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 35(5), 1299-1312.