Intelligent System for Radiological Diagnostic Assistance with an Explainable Interface

Application to Multi-Pathology Detection on Chest Radiographs

Chahnez Cherif

Master of Science in Computer Science
Faculty of Sciences and Techniques of Sidi Bouzid (FSTSB)
University of Kairouan

Proposed and Supervised Project by:

Prof. Hela Ltifi

December 2025

Abstract

This project proposes an intelligent system for assisting in thoracic radiological diagnosis. It addresses the increasing volume of examinations and multi-pathology complexity by combining accuracy and explainability. Based on the ChestX-ray14 (NIH) dataset (112,120 images, 15 pathologies), ConvNeXt-Tiny is fine-tuned to achieve a mean AUC of 0.7368. Grad-CAM++ provides visual explainability through activation maps overlaid on images, highlighting discriminative anatomical regions. A Streamlit interface enables image upload, probabilistic predictions, and interactive visualization of explanations. Results show effective detection with coherent localization, enhancing clinical trust. This transparent system offers a valuable tool for radiological decision support.

Keywords: Radiological diagnostic aid, Medical artificial intelligence, ConvNeXt, Explainability, Grad-CAM++, ChestX-ray14

1 Introduction

In modern healthcare, chest X-ray analysis faces major challenges due to increasing examination volumes and multi-pathology complexity. Artificial intelligence-based decision support systems (DSS) offer valuable assistance to radiologists, but clinical adoption is limited by the opacity of deep learning models ("black boxes").

This project proposes a multi-level framework that optimizes both accuracy and interpretability for radiological decision support. Built on the ChestX-ray14 (NIH) dataset (112,120 images, 15 pathologies), it uses fine-tuned ConvNeXt-Tiny (mean AUC 0.7368) and Grad-CAM++ for visual explainability. A Streamlit interface allows image loading, probabilistic predictions, and activation map visualization.

The dual objective is to improve diagnostic performance while providing transparency for radiologist validation, meeting essential clinical requirements for trust and justification.

The article is organized as follows: Section 2 presents the theoretical foundations, Section 3 the proposed framework, Section 4 implementation and results, Section 5 discussion, and Section 6 conclusion.

2 Theoretical background

2.1 Radiology-based medical decision support systems

Radiological decision support systems (DSS) analyze medical images to aid clinical diagnosis and treatment decisions [8]. Chest X-rays, the most common radiological examination, pose challenges due to high volume and multi-pathology complexity [14]. Machine learning models automate detection, improving efficiency and reducing inter-observer variability [7].

However, clinical adoption remains limited by the "black-box" nature of deep models, which lack transparency in their decision-making process [10]. Radiologists require understandable explanations to validate AI recommendations, a critical factor in high-stakes medical contexts [2].

Recent studies (2019-2023) emphasize the need for hybrid systems that combine precisionsion with interpretability to facilitate clinical integration [11].

2.2 ConvNeXt architecture

ConvNeXt [6] represents a modernization of convolutional networks, incorporating transformer-inspired designs (larger kernels, layer normalization, GELU activations) while preserving CNN efficiency. This architecture achieves competitive performance with reduced computational requirements, making it suitable for medical applications with resource constraints.

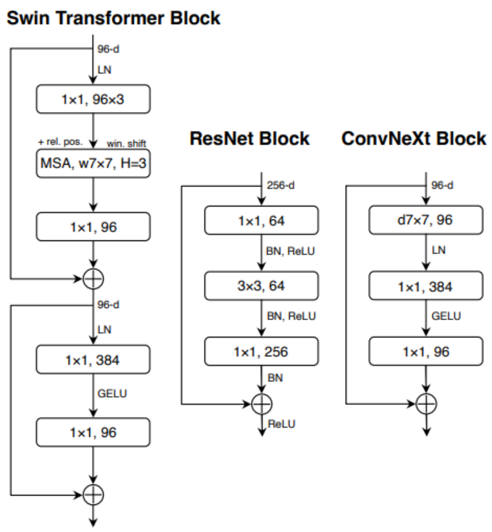

Figure 1: ConvNeXt block design and comparison with ResNet and Swin Transformer [6].

Applications in medical imaging (2023-2025) demonstrate ConvNeXt's effectiveness in classification and segmentation tasks [15, 9].

2.3 Explainable artificial intelligence

Explainability is essential in medical AI to build clinical trust [12]. Post-hoc techniques provide information about model decisions. Grad-CAM++ [1] improves on Grad-CAM by positively weighting gradients, offering a more precise localization of discriminative regions.

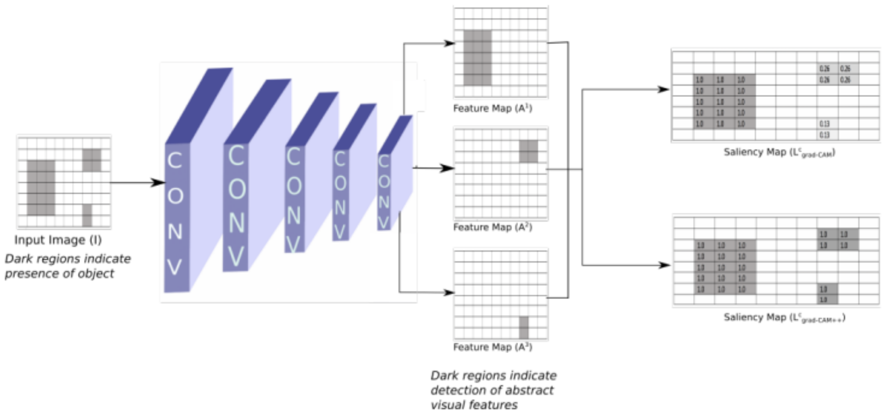

Figure 2: Examples of Grad-CAM++ visualizations on chest X-rays highlighting pathological regions [1].

Recent studies (2017-2021) apply Grad-CAM++ to chest radiograph analysis for pathological location and clinician validation [13, 5].

3 Proposed Multi-Level Framework for Radiological DSS

The proposed framework adopts a hierarchical approach with four interconnected modules, enabling robust and explainable analysis of chest radiographs (Figure 3).

3.1 Module 1: Image Preprocessing

This module forms the basis of the pipeline. It includes adaptive resizing to 224×224 pixels, normalization according to ImageNet statistics, intensity correction (CLAHE if necessary), and medical augmentations that preserve anatomical features (moderate rotations, horizontal flips). These steps ensure a consistent and robust image representation for the model.

3.2 Module 2: Classification with ConvNeXt

ConvNeXt-Tiny is fine-tuned for multi-label classification (15 classes). The training uses BCEWithLogitsLoss, the AdamW optimizer, and a CosineAnnealingWarmRestarts scheduler. The performance achieved reaches an average AUC of 0.7368 in validation (split

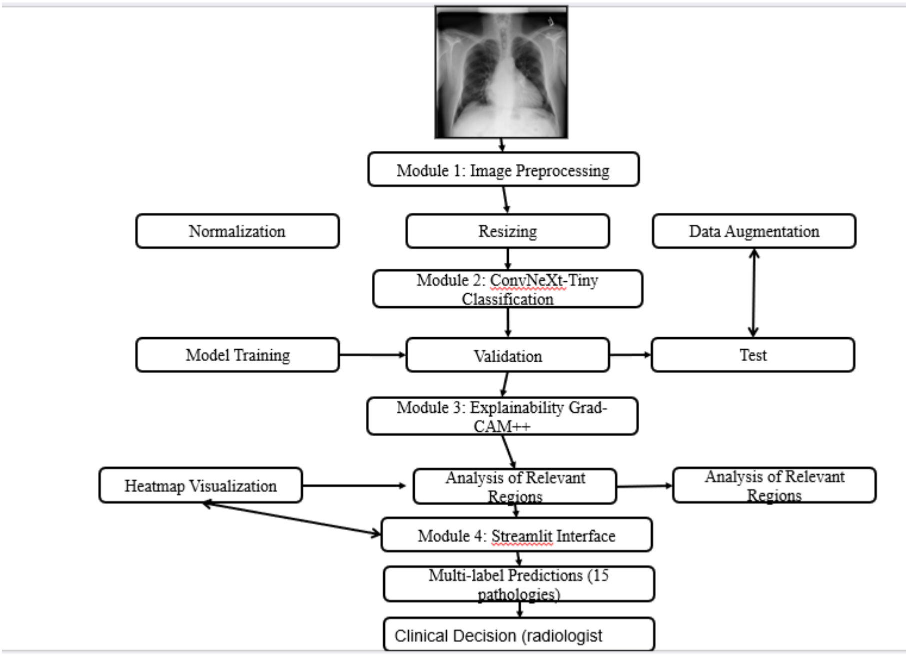

Figure 3: Multi-level framework of the radiological diagnostic support system.

patient-wise), demonstrating reliable detection of common pathologies (cardiomegaly, pneumonia, pleural effusion).

3.3 Module 3: Explainability with Grad-CAM++

Grad-CAM++ is implemented to generate activation maps from the last normative layer of the model. These maps are superimposed on the original radiograph with 40‘% opacity, clearly highlighting the anatomical regions most influential in the model's decision-making process. This post-hoc technique provides precise localization of discriminating features, enabling intuitive clinical interpretation and validation of the system's reasoning.

3.4 Module 4: Decision-Making Interface

The interface, developed with Streamlit, offers a smooth and intuitive clinical workflow. Users can load a chest X-ray, view color-coded probabilistic predictions based on risk level, and interactively generate Grad-CAM++ visualizations. The seamless integration of these four modules results in a coherent end-to-end system that supports radiologists by combining diagnostic performance with complete transparency.

4 Case Study: Multi-Pathology Detection on ChestX-ray14

4.1 Dataset

The primary dataset used in this study is ChestX-ray14 [8], a widely adopted benchmark for multi-label classification in chest imaging. It contains 112,120 annotated frontal chest radiographs labeled with 15 thoracic pathologies, including cardiomegaly, pneumonia, atelectasis, and pneumothorax 4. Each image is accompanied by rich metadata , facilitating in-depth analysis.

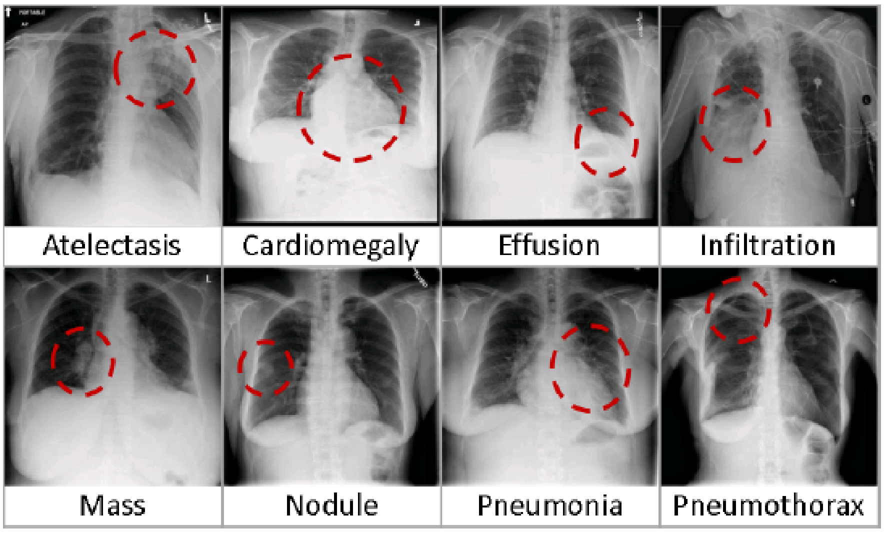

Figure 4: Eight common thoracic pathologies observed on chest radiographs illustrate the complexity of the automated diagnostic task. [8].

To assess the generalizability of our approach, complementary datasets such as MIMIC-CXR [4] and CheXpert [3] were also considered. 2).

4.2 Results

The fine-tuned ConvNeXt model achieves a mean macro AUC of 0.7368 in validation (split patient-wise). The interface allows for complete interactive analysis.

Grad-CAM++ visualizations show consistent localization of relevant anatomical anomalies, validating the model's reasoning and enhancing clinical confidence.

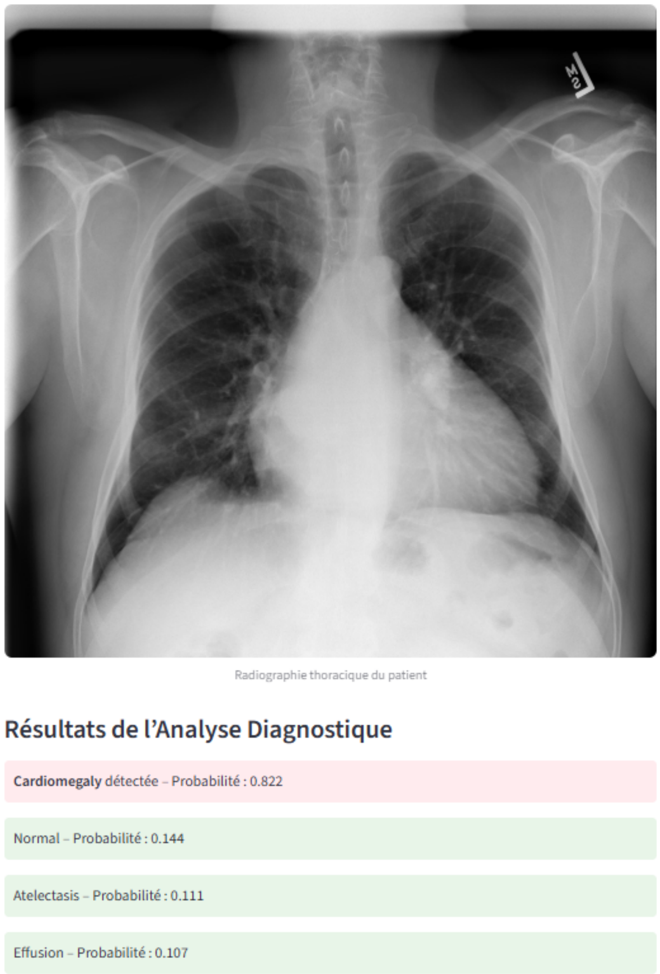

Figure 5: Streamlit interface when loading a chest X-ray and displaying probabilistic predictions .

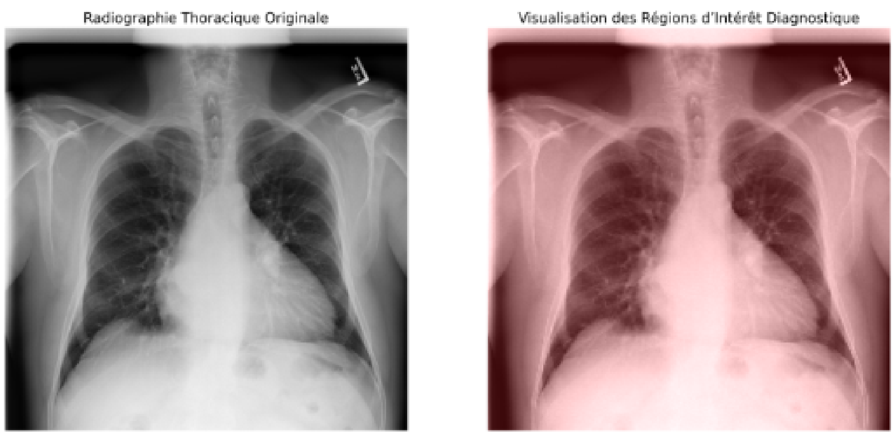

Figure 6: Grad-CAM++ visualization after button activation .

5 Discussion

The results obtained (mean AUC 0.7368) are satisfactory considering the hardware constraints (CPU training) and the labeling noise of ChestX-ray14. The model effectively detects common pathologies.
Explainability via Grad-CAM++ is the main advantage: heatmaps systematically locate abnormalities in the relevant anatomical regions (heart for cardiomegaly, lungs for pneumonia). This consistency strengthens clinical confidence.
The Streamlit interface offers an intuitive workflow closely aligned with radiological practice.
However, performance is lower than more resource-intensive models; there is no clinical validation or multimodal integration, and the model is not fast.
Extending the model to the MIMIC-CXR dataset, deploying it as a web application, and conducting clinical trials are the main directions for this work to improve its effectiveness.

6 Conclusion

This project proposes an intelligent system for assisting with thoracic radiological diagnosis, combining performance, explainability, and ergonomics. Based on ConvNeXt-Tiny and Grad-CAM++, it achieves an average AUC of 0.7368 on ChestX-ray14 while providing consistent visualizations of discriminating regions. The Streamlit interface ensures an intuitive clinical workflow, fostering confidence and adoption by radiologists. This system helps facilitate the reporting process and prevent medical errors by supporting clinical decision-making.

References

[1] Aditya Chattopadhyay et al. “Grad-CAM++: Improved Visual Explanations for Deep Convolutional Networks”. In: *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2018, pp. 839–847. URL: https://arxiv.org/abs/1710.11063.

[2] Andreas Holzinger, Georg Langs, Helmut Denk, et al. “What Do We Need to Build Explainable AI Systems for the Medical Domain?” In: *arXiv preprint arXiv:1712.09923* (2019). URL: https://arxiv.org/abs/1712.09923.

[3] Jeremy Irvin et al. “CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison”. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 33.01 (2019), pp. 590–597. URL: https://arxiv.org/abs/1901.07031.

[4] Alistair E. W. Johnson et al. “MIMIC-CXR: A large publicly available database of labeled chest radiographs”. In: *arXiv preprint arXiv:1901.07042* (2019). URL: https://arxiv.org/abs/1901.07042.

[5] Paras Lakhani and Baskaran Sundaram. “Deep Convolutional Neural Networks for End-to-End Tuberculosis Detection from Chest Radiographs”. In: *Radiology* 284.2 (2017), pp. 574–582. DOI: 10.1148/radiol.2017162326. URL: https://pubs.rsna.org/doi/10.1148/radiol.2017162326.

[6] Zhuang Liu et al. “A ConvNet for the 2020s”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022. URL: https://arxiv.org/abs/2201.03545.

[7] Hieu Pham, Yuchen Qiu, and Hongyuan Zha. “Interpretable Machine Learning for Medical Image Analysis: A Survey”. In: *arXiv preprint arXiv:2103.02401* (2021). URL: https://arxiv.org/abs/2103.02401.

[8] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, et al. “CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning”. In: *arXiv preprint arXiv:1711.05225* (2017). URL: https://arxiv.org/abs/1711.05225.

[9] Saikat Roy et al. “MedNeXt: Transformer-Driven Scaling of ConvNets for Medical Image Segmentation”. In: *arXiv preprint* arXiv:2303.09975 (2023). URL: https://arxiv.org/abs/2303.09975.

[10] Cynthia Rudin. “Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead”. In: *Nature Machine Intelligence* 1.5 (2019), pp. 206–215. DOI: 10.1038/s42256-019-0048-x. URL: https://www.nature.com/articles/s42256-019-0048-x.

[11] Erico Tjoa and Cuntai Guan. “A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI”. In: *IEEE Transactions on Neural Networks and Learning Systems* (2020). DOI: 10.1109/TNNLS.2020.3027314. URL: https://ieeexplore.ieee.org/document/9254674.

[12] Eric J. Topol. “High-performance medicine: the convergence of human and artificial intelligence”. In: *Nature Medicine* 25.1 (2019), pp. 44–56. DOI: 10.1038/s41591-018-0300-7. URL: https://www.nature.com/articles/s41591-018-0300-7.

[13] Sheng Wang, Bo Kang, Jinlu Ma, et al. “Chest X-ray Abnormality Detection Using Deep Learning with Localization”. In: *IEEE Access* 9 (2021), pp. 35679–35689. DOI: 10.1109/ACCESS.2021.3062016. URL: https://ieeexplore.ieee.org/document/9442024.

[14] Xiaosong Wang et al. “ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases”. In: *arXiv preprint arXiv:1705.02315* (2017). URL: https://arxiv.org/abs/1705.02315.

[15] Xuejun Zhang et al. “Wavelet-Guided Multi-Scale ConvNeXt for Unsupervised Medical Image Registration”. In: *Bioengineering* 12.4 (2025), p. 406. DOI: 10.3390/bioengineering12040406. URL: https://www.mdpi.com/2306-5354/12/4/406.