Explainable Multi-Label Chest X-ray Analysis for Severity-Aware Triage

Authors: Paulius Bundza; Assoc. Prof. Dr. Justas Trinkūnas (VILNIUS TECH, Dept. of Information Systems). paulius.bundza@stud.vilniustech.lt, justas.trinkunas@vilniustech.lt

VILNIUS TECH
Vilnius Gediminas Technical University

Introduction

- **Problem & need.** Chest X-rays are high-volume and time-critical, but interpretation is resource-intensive and error-prone amid radiologist shortages. An AI “second reader” can reduce turnaround and variability.
- **Objective.** Introduce MCADS - a multi-label deep-learning system that detects 18 chest radiographic abnormalities in a single image and explains predictions via Grad-CAM.
- **Novelty.** Goes beyond common 14-label setups (e.g., ChestX-ray14/CheXpert) to cover 18 findings, pairing broad multi-label coverage with interactive explainability and an end-to-end web app.
- **Operational value.** Embedded case-level prioritization (insignificant ≤19%, moderate 20-30%, significant ≥31% avg. probability) to surface urgent cases.

Methods

- **Data & labels.** Evaluated across 8 public datasets (NIH ChestX-ray14, CheXpert, MIMIC-CXR [CheXpert labels], Google DS1, RSNA Pneumonia, SIIM-ACR Pneumothorax, PadChest, VinDr-CXR). Primary metric: AUC-ROC per pathology.
- **Model.** DenseNet-121 (TorchXRayVision) [1] with sigmoid heads for 18 labels; pretrained on multi-dataset corpora; post-hoc temperature scaling and per-label thresholds for calibration.
- **Explainability.** Grad-CAM heatmaps for each predicted abnormality; overlays stored with probabilities for transparent review and QA.
- **System architecture & workflow.** Django (ASGI) web tier + Celery/Redis async workers; Nginx/Gunicorn deployment; media + PostgreSQL DB persistence; OOD gating via autoencoder. Clinical flow (BPMN). Technologist acquires CXR → MCADS preprocesses & infers (multi-label + Grad-CAM) → Radiologist reviews, confirms/edits → report forwarded to lung disease specialist for future treatment plan.
- **UI.** Single-page workflow: upload → progress → results with summary findings, per-label probabilities, and clickable Grad-CAM thumbnails; history & admin views included.

Results

- **Classification performance (AUC-ROC examples).** Strong, cross-dataset discrimination for many labels; examples from the summarized table: Cardiomegaly up to 0.93 (PadChest), Effusion up to 0.95 (PadChest), Pneumothorax up to 0.93 (VinDr), Pneumonia 0.84 (CheXpert). Full per-dataset/per-label grid in Table 1.
- **Generalization & usability.** Generalizes across 8 datasets without retraining, maintaining state-of-the-art DenseNet-121 accuracy when served through the web app (no performance drop due to integration). Inference latency: ~5-30 s per image on a 2-core CPU after warm-up; model caching reduces subsequent latency. Explanations align with anatomy (e.g., consolidation regions, pleural spaces), supporting clinician trust.
- **Limitations & next steps.** Susceptible to dataset bias/domain shift; thresholds may need site-specific tuning; post-hoc saliency can be misleading if misused; prospective clinical validation and multimodal fusion are future priorities.

Dataset Pathology	NIH ChestX-ray14	Google	RSNA	SIIM	PadChest	VinBrain	CheXpert	MIMIC-CXR
Atelectasis	0.76	-	-	-	0.77	0.67	0.91	0.88
Cardiomegaly	0.88	-	-	-	0.93	0.90	0.91	0.88
Consolidation	0.77	-	-	-	0.88	0.93	0.90	0.91
Edema	0.85	-	-	-	0.97	-	0.92	0.92
Effusion	0.85	-	-	-	0.95	0.87	0.94	0.92
Emphysema	0.73	-	-	-	0.87	-	-	-
Fibrosis	0.72	-	-	-	0.94	-	-	-
Hernia	0.91	-	-	-	0.96	-	-	-
Infiltration	0.68	-	-	-	0.85	0.86	-	-
Mass	0.80	-	-	-	0.85	-	-	-
Nodule	0.69	-	-	-	0.85	-	-	-
Pleural Thickening	0.74	-	-	-	0.79	0.84	-	-
Pneumonia	0.71	-	0.86	-	0.82	-	0.84	0.82
Pneumothorax	0.75	0.85	-	0.79	0.81	0.93	0.85	0.81
Lung Opacity	-	0.92	0.88	-	0.87	0.85	0.87	0.86
Fracture	-	0.74	-	-	0.74	-	0.74	0.74
Enlarged Cardiomediastinum	-	-	-	-	-	-	0.78	0.84
Lung Lesion	-	-	-	-	-	-	0.84	0.82

Table 1. DenseNet-121 AUC-ROC values for 18 pathologies on held-out test sets (20% split) [1]

System screenshots

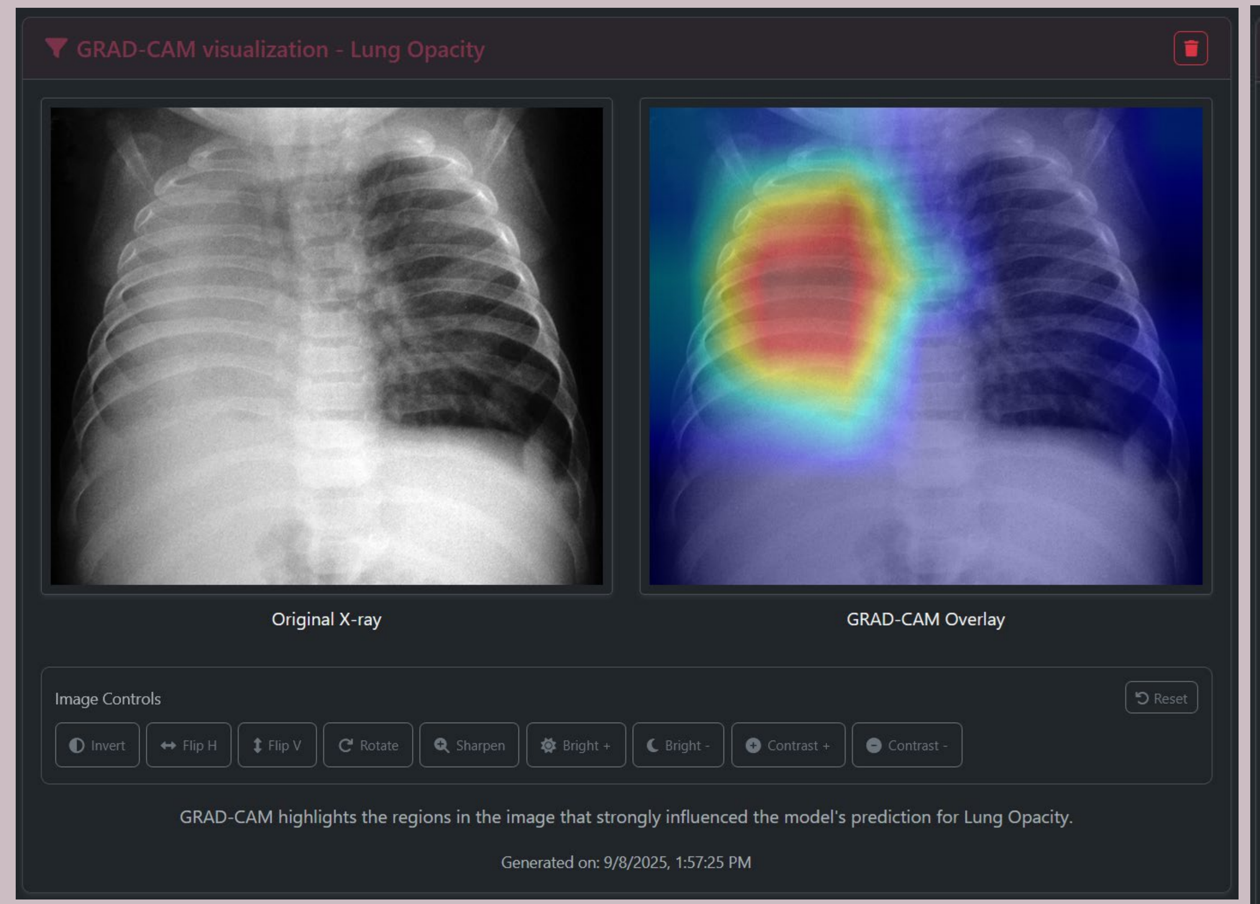


Figure 1. GUI of MCADS - Grad-CAM visualization

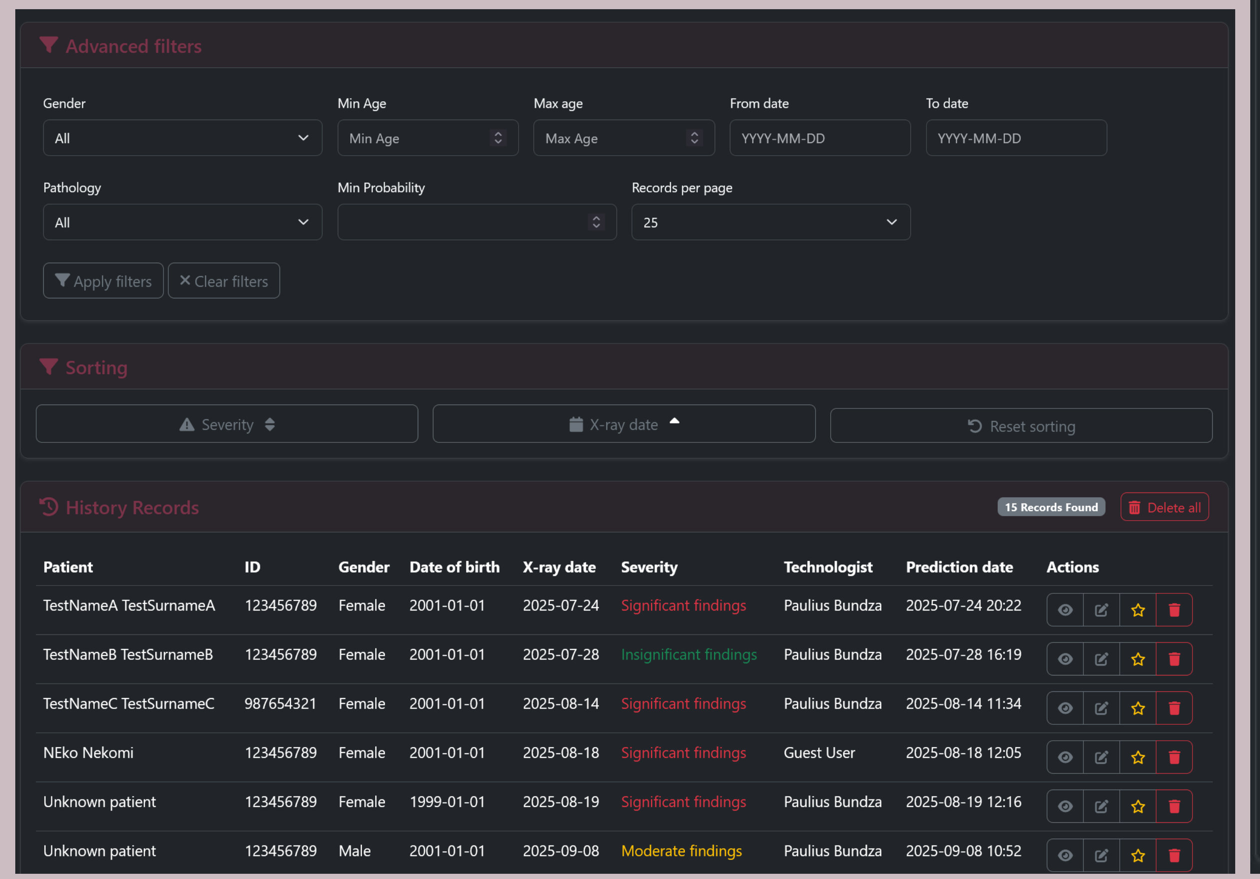


Figure 2. GUI of MCADS - prediction history

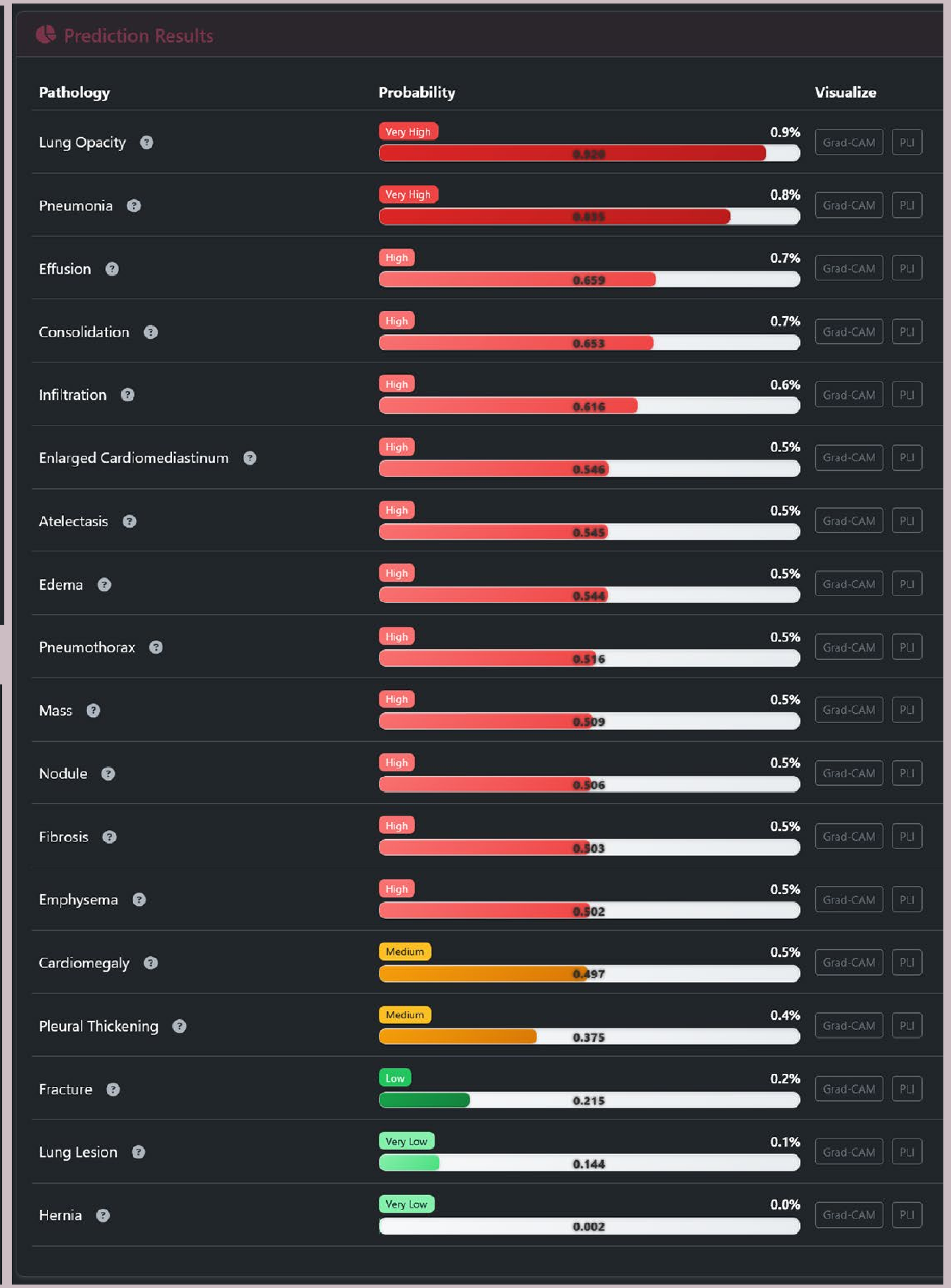


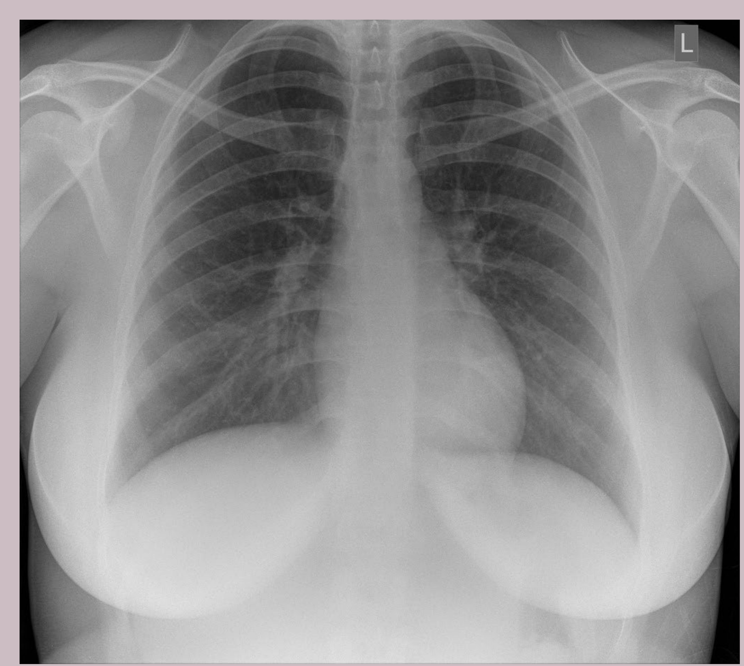
Figure 3. GUI of MCADS - bottom section of results page

References

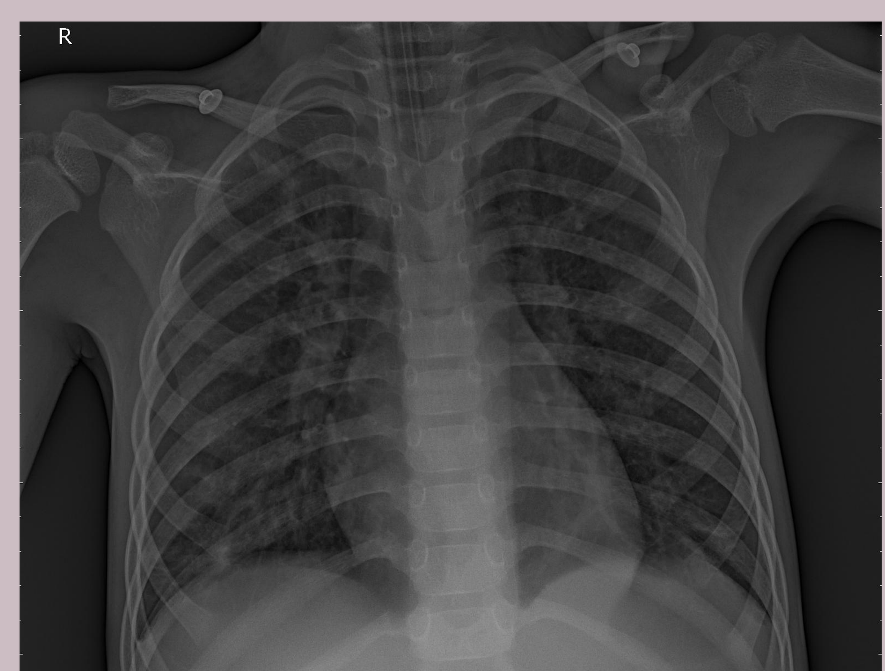
[1] J. ~P. Cohen *et al.*, “TorchXRayVision: A library of chest X-ray datasets and models,” in *Proceedings of the 5th International Conference on Medical Imaging with Deep Learning (MIDL)*, in PMLR, vol. 172. 2022, pp. 1-19. doi: 10.48550/arXiv.2111.00595.

[2] Mooney, P. (2018). Chest X-ray images (Pneumonia). https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

Live demo



Healthy patient [2]



Pneumonia + lung opacity [2]

Username: DAMSS1, Password: DAMSS2025!
Username: DAMSS3, Password: DAMSS2025!
Link to the images are in the footer
Website: https://mcads.casa/

Test these X-ray images!

Check out this system!