Proceedings of ICEGEE 2026 Symposium: AI-Based Medicine and Biological Data Analysis
DOI: 10.54254/2753-8818/2026.AU34351

Predicting Chest X-ray Pathology from Admission Data: A Cardiac Triage System for Life-saving Early Intervention

Yunhan Tong

McLean High School, Mc Lean, USA
helentong6666666@gmail.com

Abstract. Cardiovascular disease remains the leading cause of death globally, yet a critical bottleneck exists between obtaining electrocardiograms within minutes and chest X-rays in 24-72 hours, limiting life-saving early interventions. Here I present a transformer-based multimodal framework that predicts chest X-ray pathology findings from routine admission data, including 12-lead electrocardiograms and blood laboratory results, collected 24-72 hours before radiography. Using the Symile-MIMIC dataset of 11,622 hospital admissions, my approach employs multi-scale temporal transformers to capture long-range dependencies in ECG signals and cross-modal attention to integrate complementary electrical and biochemical information. The model achieved higher performance than ResNet-based baselines not only on predicting cardiac pathologies but also on mortality rate. These results demonstrate that transformer-based multimodal learning could enable earlier triage decisions, reduce unnecessary imaging, and facilitate timely interventions in critical cardiac emergencies, suggesting a more reliable pathway in both hospital and urgent care settings.

Keywords: Transformer, cardiac pathology, disease triage, chest X-ray

1. Introduction

Cardiovascular disease (CVD) remains the leading cause of death worldwide. In the United States, nearly one person dies every 34 seconds from CVDs, with 919,032 deaths reported in 2023 alone [1]. This calls for an urgent need for early detection and more timely diagnosis. However, emergency departments face a critical diagnostic bottleneck: while 12-lead electrocardiograms (ECGs) can be obtained within minutes, chest X-rays (CXRs) often require patient transport, technician availablity, and interpretation of radiologist, delaying diagnosis by 24-72 hours. This time gap limits the potential for life-saving early interventions. If cardiac electrical patterns could predict subsequent radiographic findings, clinicians could act much sooner.

Current cardiac AI systems face two fundamental challenges. First, they typically lose temporal dynamics. Traditional time-series analysis usually processes signals frame-by-frame, missing the long-range temporal relationships that define cardiac function. The heart's contractile performance emerges from coordinated chamber motion across full cardiac cycles. For instance, early diastolic filling patterns can predict systolic performance, and subtle timing abnormalities can indicate regional dysfunction [2]. Second, current systems lack cross-modal integration. Clinical practice collects ECGs and lab findings as independent assessments, yet both reflect the underlying

DOI: 10.54254/2753-8818/2026.AU34351

cardiovascular pathophysiology. Electrical changes often precede structural abnormalities by hours or even days [3], while blood biomarkers such as cholesterol and troponin indicate cardiac health [4]. Treating these assessments separately discards valuable predictive relationships between early electrical changes and subsequent structural findings.

I developed a transformer-based triage system that predicts chest X-ray findings before imaging is performed. Using only routine admission data (ECG recordings and blood tests), my model forecasts cardiac abnormalities 24-72 hours in advance with clinically meaningful accuracy. This predictive capability improves patient triage: high-risk patients receive expedited imaging and treatment, while low-risk cases avoid unnecessary radiation exposure and imaging costs. This is vital for life-saving early interventions and has the potential of translating into a clinically reliable triage routine. Beyond the hospital setting, the system can integrate with wearable ECG monitors for continuous cardiac risk assessment in daily life.

2. Materials

I used the Symile-MIMIC dataset [5, 6], a multimodal clinical dataset consisting of CXRs, ECGs, and blood laboratory tests. The dataset contains 11,622 hospital admissions from 9,573 adult patients. Each admission includes an ECG scanning and blood laboratory measurements collected within 24 hours of hospital admission, followed by a CXR collected 24-72 hours post-admission. Only the earliest available ECG, lab results, and CXR were selected per admission.

2.1. ECG

ECGs were obtained from the MIMIC-IV-ECG module. Studies with missing values or zero signals were excluded. Each ECG consists of 12 leads with a fixed length of 5,000 samples, stored as a tensor of shape ($1 \times 5000 \times 12$). Signals were normalized to range from -1 to 1 prior to modelling.

2.2. Laboratory results

Laboratory data were limited to the 50 most common blood tests recorded within 24 hours of admission, spanning hematologic, metabolic, coagulation, and blood gas measurements. For modelling, labs were represented using a 100-dimensional vector:

- the first 50 dimensions contain lab values transformed into percentiles using the empirical cumulative distribution function (CDF) computed on the training set,
- the remaining 50 dimensions are binary indicators denoting whether each lab value was observed.

When a lab value was missing, it was imputed using the mean percentile of that lab computed from the training data.

2.3. CXRs

Although CXRs were not used as predictive inputs in my cardiac disease modelling task, they are part of the original dataset construction. CXRs include labels for multiple pathological conditions (e.g., cardiomegaly and edema), encoded as 1 for positive, 0 for negative, and minus 1 for uncertain findings. Consistent with prior work, uncertain labels were imputed as negative.

DOI: 10.54254/2753-8818/2026.AU34351

3. Methods

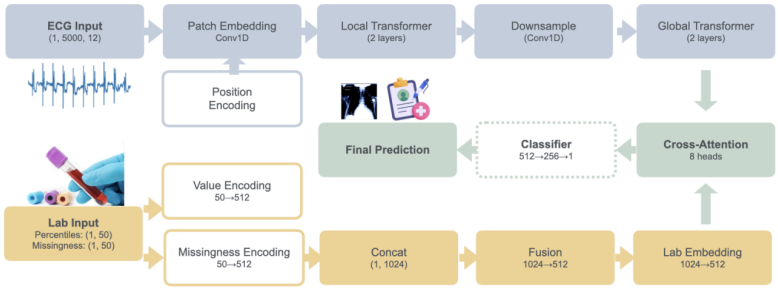

Figure 1. Multimodal transformer architecture for cardiac triage. ECG signals (12-lead, 5000 timesteps) are processed through a multi-scale temporal transformer with local and global attention layers. Laboratory values (50 tests) are encoded via dual pathways for observed values and missingness indicators. Cross-modal attention fuses ECG and lab embeddings to predict chest X-ray pathology and mortality rate

I used transformers (Fig. 1) as the backbone of my system. Transformers offer unique advantages for temporal medical data analysis and cross-modal prediction [7, 8]. Specifically, the attention mechanisms of transformers can model long-range temporal dependencies within cardiac signals, relating events separated by entire cardiac cycles or extended time periods. The same attention-based approach can capture cross-modal relationships, learning how patterns in one measurement type (e.g., ECG ST-segment changes) predict findings in another (e.g., pulmonary edema on chest imaging). Unlike convolutional neural networks (CNN) with limited receptive fields, transformers process the entire ECG sequence, identifying subtle temporal patterns that may indicate developing pathology.

3.1. ECG encoder: multi-scale temporal transformer

ECG signals were processed using a multi-scale transformer architecture. Raw 12-lead ECG data (5000 timesteps) were first divided into non-overlapping patches of 50 timesteps, reducing the sequence length to 100 patches. Each patch was projected to a 512-dimensional embedding space and augmented with learned positional encodings to preserve temporal order. The architecture consisted of two transformer stages to capture patterns at different timescales. Local transformer layers (2 layers, 8 attention heads) first modelled beat-level patterns within the sequence. The representation was then downsampled by a factor of 2 using strided convolution, reducing computational cost while retaining salient features. Global transformer layers (2 layers, 8 attention heads) subsequently captured long-range dependencies across the entire ECG, such as rhythm variations and sustained ST-segment changes. Each transformer layer used 4× expansion in the feed-forward network with GELU activation [9] and pre-layer normalization for training stability.

3.2. Laboratory encoder: dual pathway processing

Laboratory values were encoded through separate pathways for observed percentile values and missingness indicators. The 50 lab percentile values were processed through a two-layer feed-forward network (50→256→512 dimensions). The 50 binary missingness indicators followed an identical architecture. This dual-pathway design allowed the model to distinguish between low lab values and missing measurements, which is an important distinction in clinical settings. The value

DOI: 10.54254/2753-8818/2026.AU34351

and missingness representations were concatenated and fused through a final projection layer with layer normalization and dropout (p=0.2), producing a 512-dimensional lab embedding.

3.3. Multimodal fusion

The ECG and lab embeddings were integrated via bidirectional cross-attention [7]. ECG features attended to lab embeddings as context, allowing laboratory biomarkers to guide interpretation of electrophysiological patterns. Similarly, lab features attended to the ECG temporal sequence, enabling cardiac electrical activity to inform biochemical markers. Each cross-attention operation used 8 attention heads with residual connections to preserve modality-specific information. The cross-attended representations were averaged to produce a unified 512-dimensional embedding. This final embedding was then passed to a classifier to predict cardiac pathologies and morality rate.

3.4. Baseline method: ResNet-10

I used ResNet10 [10] as a baseline benchmark for the transformer approach. Specifically, the ECGs were projected into an embedding space via a ResNet10 network trained by Simple Model-Agnostic Learning (Symile) [6]. The ECG embeddings were subsequently combined with the Labs embedding and fed into a multilayer perceptron (MLP) to give final predictions.

3.5. Data splits

The dataset was divided into training and test sets at the patient level to prevent data leakage from repeated admissions. The training set contained 10,000 admissions from 8,374 patients, and the test set contained 750 admissions from 738 patients. Patient identities were mutually exclusive between splits, ensuring no information leakage from shared patient histories. Laboratory percentile transformations were computed exclusively on the training set and applied to the test set. All reported performance metrics were evaluated on the held-out test set in a single final evaluation.

3.6. Optimization

All models were trained using binary cross-entropy loss with logits. I used the AdamW optimizer with an initial learning rate of 0.0001 and weight decay of 0.1. Learning rate was reduced by a factor of 0.5 when validation performance plateaued for 3 consecutive epochs, using ReduceLROnPlateau scheduling based on AUROC. Models were trained for 100 epochs with a batch size of 32. Training was performed on a single NVIDIA RTX 4090 GPU using PyTorch [11].

3.7. Evaluation metric

I used Area Under the Receuver Operating Characteristic Curve (AUROC) scoreto evalute the model and baseline. AUROC plots the True Positive Rate against the False Positive Rate, showing how well the model ranks positive instances higher than negative ones. It represents how well the model ranks positive instances higher than negative ones, with 1.0 being perfect and 0.5 being random chance.

DOI: 10.54254/2753-8818/2026.AU34351

4. Experiments

4.1. Cardiac pathologies prediction

I first evaluated prediction accuracy for two cardiac pathologies, Cardiomegaly and Pulmonary Edema, using both the transformer and baseline approaches. Given that the input data includes only ECGs and laboratory results, I anticipated moderate Area Under the Curve (AUC) scores, as these conditions are conventionally diagnosed through chest radiography rather than electrophysiological recordings. Cardiomegaly indicates abnormal enlargement of one or more cardiac chambers [12], while pulmonary edema reflects excessive fluid in the lungs as a result of heart dysfunction [13]. Both pathologies are typically defined chest X-rays and remain hidden on ECGs in traditional methods. Detecting these pathologies from ECGs before chest X-rays are taken will therefore have strong clinical implications.

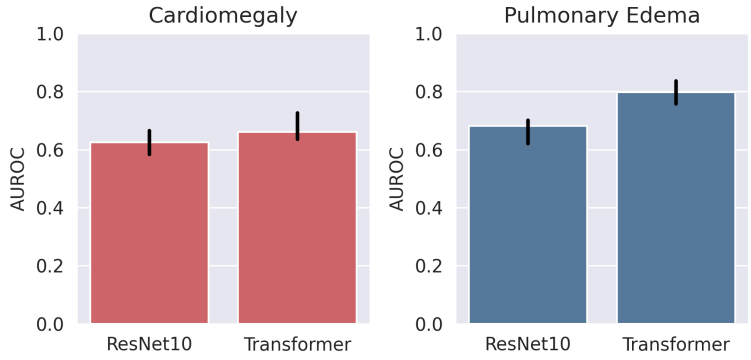

Figure 2. Pathology prediction performance. AUROC scores for Cardiomegaly and Pulmonary Edema prediction on the test set (n=750). Error bars indicate 95% confidence intervals. The transformer outperformed ResNet-10 baseline on both pathologies, with the largest improvement observed for Pulmonary Edema (0.80 vs 0.68)

For both pathologies, my proposed transformer method (Fig. 2) yielded higher AUC scores. On Cardiomegaly prediction, transformer gave an AUC score of 0.6605 (95% confidence interval: [0.62 , 0.70]). On Pulmonary Edema, transformer gave an AUC score of 0.80 (95% confidence interval: [0.76, 0.84]), significantly higher than the AUC score 0.68 given by ResNet10 (95% confidence interval: [0.64 , 0.73]).

4.2. Mortality rate prediction

I next aim to predict mortality rate using the proposed transformer (Fig 3). ECG is typically a good predictor of death rate, particularly for sudden cardiac death and cardiovascular mortality. While standard ECG parameters have already been linked to increased risk of cardiovascular death, advanced AI models can extract more subtle information, further improving mortality prediction capabilities, even in cases a physician might interpret as normal. Transformer demonstrated higher AUROC than ResNet10, suggesting it better captures the subtle information in ECGs in predicting mortality rate.

DOI: 10.54254/2753-8818/2026.AU34351

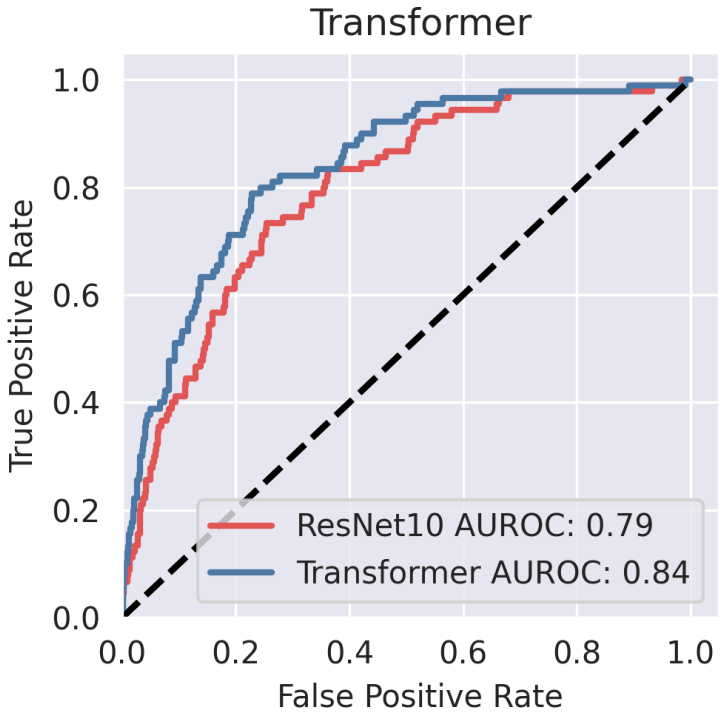

Figure 3. Mortality prediction performance. ROC curves comparing transformer and ResNet-10 baseline for in-hospital mortality prediction. The transformer achieved higher AUROC, indicating improved capture of prognostic information from ECG and laboratory inputs

5. Conclusion

In this project, I proposed a transformer-based multimodal framework that leverages routinely available ECG recordings and blood laboratory results to predict pathologies typically identified on chest X-rays and patient mortality risk. By modelling long-range temporal dynamics in ECG signals and integrating complementary biochemical information through cross-modal attention, my approach addresses two key limitations of prior cardiac AI systems: loss of temporal context and lack of multimodal integration. Evaluated on the Symile-MIMIC dataset, the proposed model consistently demonstrated better AUROCs than a ResNet-based baseline on Cardiomegaly, Pulmonary Edema and mortality prediction. These findings support the hypothesis that early electrical and biochemical signatures can anticipate downstream structural abnormalities later observed on chest X-rays and capture subtle prognostic electrophysiological patterns that may not be apparent through standard clinical interpretation.

I specifically focused on Cardiomegaly and Pulmonary Edema because these conditions are conventionally defined by chest X-ray findings collected 24-72 hours after admission and are often not directly detectable from ECGs using traditional methods. While these two pathologies served as important clinical cases, the proposed framework is not limited to them and can be readily extended to other cardiac conditions or continuous clinical outcomes.

The proposed approach has many potential applications. First, by enabling earlier identification of high-risk patients from routine admission data, it holds the potential to assist clinicians in making more effective triage decisions. This could, in turn, lead to a reduction in unnecessary imaging procedures for low-risk patients, while ensuring timely intervention for those at high risk. This has

DOI: 10.54254/2753-8818/2026.AU34351

the potential to lower healthcare costs, decrease patient discomfort, and ultimately improve clinical outcomes and survival rates. However, translating these predictive capabilities into safe and routine clinical workflow changes-specifically regarding the reduction of imaging-will require rigorous validation through future prospective clinical studies. More broadly, my work demonstrates the potential of transformer-based multimodal learning as a reliable clinical routine in both hospital and urgent care settings.

References

[1] C. CDC., "National center for health statistics mortality data on CDC WONDER." 2023.
[2] P. M. Mottram and T. H. Marwick, "Assessment of diastolic function: What the general cardiologist needs to know, " Heart, vol. 91, no. 5, pp. 681–695, 2005.
[3] S. M. van der Voorn, A. S. Te Riele, C. Basso, H. Calkins, C. A. Remme, and T. A. van Veen, "Arrhythmogenic cardiomyopathy: Pathogenesis, pro-arrhythmic remodelling, and novel approaches for risk stratification and therapy, " Cardiovascular research, vol. 116, no. 9, pp. 1571–1584, 2020.
[4] M. Clinic, "Understand blood tests for heart disease." Mayo Clinic, Dec. 2023. Available: https: //www.mayoclinic.org/diseases-conditions/heart-disease/in-depth/heart-disease/art-20049357
[5] A. Johnson, L. Bulgarelli, T. Pollard, S. Horng, L. Celi, and R. Mark, "MIMIC-IV (version 2.2). PhysioNet. RRID: SCR_007345." 2023.
[6] A. Saporta, A. M. Puli, M. Goldstein, and R. Ranganath, "Contrasting with symile: Simple model-agnostic representation learning for unlimited modalities, " Advances in Neural Information Processing Systems, vol. 37, pp. 56919–56957, 2024.
[7] A. Vaswani et al., "Attention is all you need, " Advances in neural information processing systems, vol. 30, 2017.
[8] F. Shamshad et al., "Transformers in medical imaging: A survey, " Medical Image Analysis, vol. 88, p. 102802, 2023, doi: https: //doi.org/10.1016/j.media.2023.102802.
[9] D. Hendrycks, "Gaussian error linear units (gelus), " arXiv preprint arXiv: 1606.08415, 2016.
[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition, " in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[11] A. Paszke et al., "Pytorch: An imperative style, high-performance deep learning library, " Advances in neural information processing systems, vol. 32, 2019.
[12] "Enlarged Heart (Cardiomegaly): What It Is, Symptoms & Treatment — my.clevelandclinic.org." https: //my.clevelandclinic.org/health/diseases/21490-enlarged-heart-cardiomegaly.
[13] "Pulmonary Edema: Causes, Symptoms, Diagnosis & Treatment — my.clevelandclinic.org." https: //my.clevelandclinic.org/health/diseases/24218-pulmonary-edema.