5. Piscopo, M. L. Solving differential equations with neural networks: Applications to the calculation of cosmological phase transitions / M. L. Piscopo, M. Spannowsky, P. Waite // Physical Review D. - 2019. - Vol. 100, no. 1. - P. 016002.

6. Nguyen, T. D. Numerical method for solving Fredholm and Volterra integral equations using artificial neural networks / T. D. Nguyen, I. Z. Akhmetov, A. F. Galimyanov // Chebyshevskii sbornik. - 2024, Vol. 25, No. 5, Pp. 2-14.

7. Габдулхаев Б.Г. Прямые методы решения сингулярных интегральных уравнений первого рода. -Казань: Изд-во Казанск. Ун-та. – 1994. 288 с.

8. Огородников, Е. Математические модели нелинейной вязкоупругости с операторами дробного интегро-дифференцирования / Е. Огородников, В. Радченко, Л. Унгарова // Вестник Пермского национального исследовательского политехнического университета. Механика. - 2018. - №. 2. - С. 147-161.

9. Unal, E. Solution of conformable fractional ordinary differential equations via differential transform method / E. Unal, A. Gokdogan // Optik. - 2017. - Vol. 128. - Pp. 264-273.

10. Allahviranloo, T. An application of artificial neural networks for solving fractional higher-order linear integro-differential equations / T. Allahviranloo, et al //Boundary Value Problems. - 2023. - T. 2023. - №. 1. - C. 1-14.

11. Gao, F. Solving fractional differential equations by using triangle neural network / F. Gao, Y. Dong, C. Chi // Journal of Function Spaces. - 2021. - Vol. 2021. - Pp. 1-7.

EXPLAINABLE MULTIMODAL AI IN CARDIOVASCULAR DISEASE PREDICTION

Umme Sania
Assistant professor, Sambhram University, Jizzakh, Uzbekistan
Usania90@gmail.com
Alishev Sherkuzi
PhD, Head of Department, Sambhram University, Jizzakh, Uzbekistan
sheralishev1985@gmail.com

Abstract: Cardiovascular disease (CVD) is the leading global cause of mortality, and prediction is difficult in resource-limited settings. Traditional models like SCORE2 and Framingham use limited variables and offer low interpretability. We propose an Explainable Multimodal AI (EMAI) framework combining echocardiography images, ECG waveforms, and clinical data. Modality-specific encoders and a fusion transformer learn cross-modality interactions for improved prediction. Explainability is provided through SHAP for feature attribution, Grad-CAM for echo visualization, and temporal saliency for ECG interpretation. Preliminary results show higher accuracy than single-modality models. The framework enhances clinical trust and supports patient-specific cardiac risk assessment.

Keywords: explainable AI, Multimodal Deep Learning, Cardiovascular Disease, Risk Prediction, Clinical Decision Support.

1. Introduction. Cardiovascular conditions develop gradually and often remain undetected until significant structural or functional impairment has occurred. In routine clinical practice, risk assessment relies on a combination of imaging interpretation, physiological monitoring, and patient history, each offering only a partial view of the underlying cardiac status. Clinicians therefore synthesize information from multiple sources such as echocardiography, electrocardiography (ECG), and biochemical profiles to form a diagnostic judgment. However, this process is highly dependent on clinical experience, can vary across practitioners, and becomes increasingly challenging in settings with limited specialist availability.

Artificial intelligence (AI) approaches have shown promise in supporting cardiovascular evaluation by automating pattern recognition and detecting subtle abnormalities not easily visible to the human eye. For example, deep learning has enabled precise estimation of ventricular function from echocardiographic sequences and the detection of structural dysfunction from ECG waveform features alone. Despite these advances, most AI systems operate as black boxes, providing predictions without revealing the rationale behind them. In a clinical context where decisions influence treatment initiation, medication management, and surgical referrals the absence of interpretability can limit trust and prevent clinical adoption.

Furthermore, a substantial portion of existing AI models rely on single data modalities, whereas cardiac decision-making is inherently multimodal. A model that interprets only an ECG or only an echocardiogram may overlook critical contextual factors such as comorbidities, medication effects, or underlying metabolic risk. Integrating multiple data types within a single predictive framework has the potential to improve diagnostic reliability, but such models must also provide transparent, clinician-aligned explanations in order to be useful in practice.

Objective. This research aims to develop an Explainable Multimodal AI (EMAI) framework for cardiovascular disease prediction, integrating:

- Echocardiography video features
- ECG waveform characteristics
- Clinical and biochemical variables

While providing layered, clinician-aligned explanations using:

- Global and individual feature attribution (SHAP)
- Visual interpretability for imaging (Grad-CAM)
- Temporal saliency mapping for ECG waveforms

Contribution. This work addresses the performance trust gap in cardiac AI by:

1. Designing a multimodal fusion model that mirrors real clinical reasoning.
2. Embedding explainability at every stage of prediction.
3. Ensuring interpretations are clinically meaningful and verifiable.

This approach strengthens the reliability, usability, and clinical acceptability of AI-driven cardiovascular risk assessment.

2. Literature review.

2.1 Cardiovascular Risk Assessment in Clinical Practice

Clinical evaluation of cardiovascular disease typically involves synthesizing information from imaging, electrophysiological signals, laboratory values, and patient history. Standard risk scoring tools, such as the Framingham Risk Score and SCORE2, provide population-derived estimates but often perform sub optimally for individual patients and across diverse ethnic populations (D’Agostino et al., 2008; Visseren et al., 2021). These models do not leverage the rich temporal and spatial information available in ECG and echocardiography data, resulting in limited predictive sensitivity for early disease stages.

2.2 Deep Learning in Cardiovascular Diagnostics.

Advances in deep learning have enabled automated cardiac phenotype detection from raw medical signals.

- **Echocardiography:** Ouyang et al. (2020) demonstrated that convolutional neural networks can estimate left ventricular ejection fraction directly from echocardiographic video with cardiologist-level performance.
- **ECG:** Attia et al. (2019) showed that deep neural networks can infer reduced ejection fraction from ECG signals alone, suggesting that electrophysiological patterns carry latent structural information.

However, most models remain unimodal, limiting their ability to generalize across heterogeneous patient profiles.

2.3 Multimodal AI for Cardiac Risk Prediction

Multimodal models seek to capture complementary characteristics of disease:

- **Echocardiography** provides structural and functional insights.
- **ECG** captures electrophysiological disturbances.
- **Clinical variables** reflect systemic risk and comorbidities.

Recent multimodal approaches (Lee et al., 2023) have shown that integrating such modalities consistently outperforms models trained on any single modality. However, fusion strategies vary, including early, late, and cross-attention fusion, with no standard consensus on optimal architecture.

2.4 The Need for Explainability in Cardiac AI

A well-documented challenge to AI adoption in cardiology is the lack of interpretability (Topol, 2019). Clinicians require transparency to:

- Validate the clinical plausibility of predictions
- Detect model bias or error
- Integrate recommendations into treatment pathways

Explainability methods include:

- **SHAP** for feature attribution (Lundberg & Lee, 2017)
- **Grad-CAM** for visual focus maps (Selvaraju et al., 2017)
- **Temporal saliency** for waveform interpretation.

However, Adebayo et al. (2018) demonstrated that many visual explanations fail sanity checks**,** emphasizing the need for careful evaluation of interpretability methods.

Literature Gap

Most existing cardiac AI models are:

1. Unimodal
2. Opaque
3. Not designed for clinician reasoning workflows.

This creates a clear need for models that are both multimodal and inherently explainable**.**

3. Methodology

3.1 Overview

This study proposes an **Explainable Multimodal AI (EMAI)** framework that integrates:

1. **Echocardiographic video sequences**
2. **12-lead ECG time-series**
3. **Structured clinical data** (e.g., age, comorbidities, lab measures)

The system outputs both **predicted cardiovascular risk** and **explanations** detailing how each modality contributed.

3.2 Data Pipeline

Step	Description
Data Acquisition	Echo, ECG, and clinical data from hospital cardiac imaging systems & electronic medical records
Preprocessing	Video normalization, ECG filtering, missing value imputation for clinical tables
Feature Encoding	Separate encoders for each modality
Multimodal Fusion	Cross-modality attention-based transformer
Prediction	Binary or continuous outcome (e.g., CVD risk score, LV dysfunction likelihood)
Explainability	SHAP, Grad-CAM, Temporal Saliency

Table-1: Multimodal Cardiovascular Risk Prediction Pipeline

3.3 Modality Encoders

3.3.1 Echocardiography Branch (Visual): A **3D Convolutional Neural Network** (e.g., R(2+1)D or ResNet-3D) extracts temporal-spatial representations.

$$h_{echo} = f_{3D-CNN}(V)$$

where V is the video tensor.

3.3.2 ECG Branch (Waveform): A **1D CNN + Transformer encoder** captures local wave morphology and long-range temporal dependencies.

$$h_{ecg} = f_{1D-CNN+Transformer}(S)$$

where S is the ECG sequence.

3.3.3 Clinical Data Branch (Structured): A **Multi-Layer Perceptron (MLP)** processes clinical variables.

$$h_{clin} = f_{MLP}(C)$$

3.4 Multimodal Fusion Layer

The three embeddings are fused using cross-attention:

$$h_{fusion} = \text{Transformer}(h_{echo}, h_{ecg}, h_{clin})$$

This allows each modality to inform and refine the others.

3.5 Prediction Layer

$$\hat{y} = \sigma(W h_{fusion} + b)$$

where σ is a sigmoid activation for risk probability.

3.6 Explainability Strategy

Explainability Method	Applied To	Output
SHAP	Clinical + Combined Prediction	Global and patient-level feature impact

Grad-CAM	Echocardiography frames	Visual activation regions
Temporal Saliency Maps	ECG waveforms	Important waveform intervals

Table-2: *Explainability Methods and Their Applications in the Multimodal Cardiovascular Risk Prediction Framework*

These explanations are designed to mirror clinical reasoning, enabling physicians to validate model decisions.

3.7 Evaluation

Model performance will be evaluated using:

- AUROC, Precision-Recall AUC
- Calibration (Brier score, ECE)
- Clinician interpretability scoring (Likert-scale usefulness and trust ratings)

External validation will be performed to assess generalizability.

4. Conceptual outcomes and expected system behavior

The proposed Explainable Multimodal AI framework integrates echocardiographic imaging, ECG waveforms, and clinical variables into a unified model for cardiovascular risk prediction. By combining structural, electrical, and patient-specific data, it offers more accurate and context-aware assessments than single-modality approaches. The framework also prioritizes interpretability through SHAP for clinical feature attribution, Grad-CAM for echocardiographic visualization, and temporal saliency for ECG reasoning. This ensures that clinicians can validate and trust model outputs. Overall, the contribution lies in: (1) a multimodal fusion architecture aligned with real clinical decision-making, and (2) an explainability strategy providing clear, clinically meaningful insights at both global and individual levels, supporting future deployment in routine cardiac care.

5. Discussion. The Explainable Multimodal AI framework presented in this work emphasizes that effective cardiovascular risk prediction requires a perspective that reflects the integrated nature of clinical decision-making. Cardiologists do not rely on a single measurement or modality; instead, they synthesize structural, electrical, and clinical context to assess disease severity and progression. The proposed architecture mirrors this reasoning style by combining echocardiographic dynamics, ECG waveform patterns, and patient-specific clinical profiles within a unified predictive model.

A distinguishing aspect of this framework is the central role of explainability. Rather than generating predictions in isolation, the system provides layered interpretability outputs that clarify how different data elements contribute to the risk estimate. This is clinically significant for three reasons:

1. **Transparency supports professional judgment** clinicians can confirm whether the model's rationale aligns with established pathophysiological markers.
2. **Interpretability assists communication with patients**, particularly in conveying why a risk is elevated and how it may be modified.
3. **Explainability functions as a safeguard**, enabling review of outlier or unexpected outputs to prevent inappropriate clinical action.

By grounding AI evaluation in clinical usability, reasoning alignment, and interpretability, the framework moves beyond accuracy as the sole performance criterion and aligns with emerging principles of trustworthy and human-centered medical AI.

6. Future work

The present work establishes a conceptual framework for an Explainable Multimodal AI system for cardiovascular risk prediction. The next phase of this research conducted during the PhD period will focus on implementation, validation, and real-world usability assessment.

Planned Future Work

1. **Dataset Acquisition and Integration**

- Collect echocardiography, ECG, and clinical datasets from partner hospitals.
- Ensure standardized preprocessing pipelines and annotation protocols.

2. **Model Development and Optimization**

- Train and fine-tune modality-specific encoders.
- Experiment with early, late, and cross-attention fusion strategies.

3. **Explainability Validation**

- Evaluate SHAP and Grad-CAM explanations with cardiologists.
- Conduct **sanity checks** to ensure explanation fidelity.
- Develop clinician-oriented dashboards for interactive interpretation.

4. **Clinical Evaluation and Real-World Testing**

- Perform internal and external validation to assess generalizability.

- Run **human-in-the-loop trials** to measure usefulness and trust.
- Integrate model outputs into clinical workflows for pilot deployment.

5. **Localization to Healthcare Context in Uzbekistan**

- Adapt prediction thresholds and risk factors to population-specific profiles.
- Align with available diagnostic pathways, workflows, and resource constraints.

7. Conclusion. This work presents an Explainable Multimodal AI framework designed to support cardiovascular risk prediction through the integration of echocardiographic, electrocardiographic, and clinical data. The approach reflects real-world clinical reasoning and prioritizes interpretability through structured explanation strategies, including feature attribution and visual and temporal saliency mechanisms. By focusing on both predictive insight and transparent justification, the framework addresses a key barrier to clinical acceptance of AI in cardiology: the need for systems that not only provide answers, but also clarify their reasoning. The proposed framework offers a foundation for developing trustworthy, clinically aligned decision support tools, reinforcing the role of AI as a partner in medical expertise rather than a replacement for it. Its structure is adaptable across clinical environments and can be extended based on population-specific considerations, healthcare practices, and system integration needs.

References

1. Adebayo J., Gilmer J., Muelly M., Goodfellow I., Hardt M., Kim B. Sanity checks for saliency maps // *Advances in Neural Information Processing Systems*. 2018. URL: https://arxiv.org/abs/1810.03292 (accessed: 11.11.2025).
2. Attia Z. I., Friedman P. A., Noseworthy P. A., Lopez-Jimenez F., Ladewig D. J., Satam G., Pellikka P. A., Carter R. E., Kapa S. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram // *Nature Medicine*. 2019. Vol. 25, No. 1. P. 70–74. DOI: 10.1038/s41591-018-0240-2.
3. D'Agostino R. B., Grundy S., Sullivan L. M., Wilson P. Validation of the Framingham coronary heart disease prediction scores // *JAMA*. 2008. Vol. 286, No. 2. P. 180–187. DOI: 10.1001/jama.286.2.180.
4. Lee Y. C., Jung W., Lee H. J., Han K., Kim Y. J. Multimodal deep learning for cardiovascular risk prediction from retinal images and clinical data // *NPJ Digital Medicine*. 2023. Vol. 6, No. 13. DOI: 10.1038/s41746-023-00766-3.
5. Lundberg S. M., Lee S.-I. A unified approach to interpreting model predictions // *Advances in Neural Information Processing Systems*. 2017. URL: https://arxiv.org/abs/1705.07874 (accessed: 11.11.2025).
6. Ouyang D., He B., Ghorbani A., Yuan N., Ebinger J., Langlotz C. P., Heidenreich P., Harrington R. A., Liang D. H., Ashley E. A., Zou J. Video-based AI for beat-to-beat assessment of cardiac function // *Nature*. 2020. Vol. 580. P. 252–256. DOI: 10.1038/s41586-020-2145-8.
7. Selvaraju R. R., Cogswell M., Das A., Vedantam R., Parikh D., Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization // *Proceedings of the IEEE International Conference on Computer Vision*. 2017. P. 618–626. DOI: 10.48550/arXiv.1610.02391.
8. Topol E. J. High-performance medicine: The convergence of human and artificial intelligence // *Nature Medicine*. 2019. Vol. 25. P. 44–56. DOI: 10.1038/s41591-018-0300-7.

СОВЕРШЕНСТВОВАНИЕ ЛИНГВОТЕХНИЧЕСКОЙ ПОДГОТОВКИ СПЕЦИАЛИСТОВ В УСЛОВИЯХ НАУЧНО-ТЕХНОЛОГИЧЕСКОЙ ТРАНСФОРМАЦИИ ОБЩЕСТВА

Хазраткулова Анжелика Валерьевна
доцент, Джизакский политехнический институт, г. Джизак, Республика Узбекистан,
Досалиев Канат Серикулы
PhD доцент, Южно-Казахстанский университет им. М. Ауэзова, Республика Казахстан

Аннотация: В данной работе рассматривается проблема совершенствования лингвотехнической подготовки специалистов в контексте научно-технологической трансформации сов ременного общества на примере Узбекистана. Анализируются текущие вызовы, связанные с интеграцией национальных кадров в международное профессиональное