Supplement

1 Background

The existing models primarily use machine learning techniques, and deep learning-based early fusion architecture for clinical predictions. For instance, the research study conducted by Pablo Juan-Salvadores et al. [1] focused on predicting MACE for patients who underwent coronary anatomy assessment with a long-term prognosis of 1 year. The study revealed that Random Forest (RF) outperformed other machine learning models (AUROC 0.8), achieving a significantly better performance. Similarly, Yijiang Zhou et al. and Jinwan Wang et al. have employed ML models to address MACE prediction for patients receiving PCI [2, 3]. Yijiang Zhou et al. devised an ensemble model that combines the capabilities of four ML models namely XGBoost, a light gradient boosting machine, a support vector machine (SVM), and a neural network and achieved an AUROC score of 0.74 [2]. Jinwan Wang et al. utilized parameters from EHR data obtained from a retrospective study involving 1004 patients who underwent coronary revascularization for 6-month MACE prediction and reported that the XGBoost algorithm demonstrates the highest AUROC score of 0.8599 [3]. In the case of coronary heart disease (CHD) risk prediction, Juan Jose Beunza et al. [4] reported that the neural network model and SVM model resulted in the highest AUROC score (0.71 and 0.75, respectively). However, only EHR data (i.e. data from a single modality) was considered in these studies.

The research conducted by Micah Liam Arthur Heldeweg et al. [5] focused on the 30-day MACE prediction by employing a multivariable logistic regression algorithm trained with multimodal data, including ECG readings, vital signs, and demographic information. Notably, an early fusion technique was utilized to handle the integration of information from multiple modalities effectively. Similarly, Jiandong Zhou et al. [6] investigated the prediction of stroke, atrial fibrillation (AF), cardiovascular (CV) mortality, and all-cause mortality using machine learning models by leveraging early fusion of multimodal data comprising ECG and EHR data. The findings highlighted XGBoost as the most effective algorithm, achieving an AUROC score of 0.80 for stroke prediction, 0.82 for AF prediction, 0.595 for CV mortality prediction, and 0.766 for all-cause mortality prediction. However, it should be noted that the early fusion technique generates a spare high-dimensional feature representation by combining the features from different modalities and assumes that there exists a fundamental alignment between the input features from different modalities.

In the context of post-PCI prognosis, we propose a deep learning-based joint fusion model to predict adverse cardiovascular outcomes using two data modalities (ECG captured as 2D image and EHR). Joint-fusion architecture with Convolutional Block Attention Module (CBAM) can propagate the loss directly from the prediction layer to the multimodal parallel feature extraction layers which allows optimization of focused co-learning between image and EHR, enhancing the integration of information from multiple sources.

2 Detailed list of clinical variables

Baseline Demographics		**Pre-Procedural Variables**		**Post-Procedural Complications**	
Age	66.92 ± 12.51	Cardiac arrest within 24 hours	1.86%	MI complication	3.50%
Male	70.70%	Pre-PCI LV Ejection fraction (≤40%)	88.83%	Cardiogenic shock complication	2.95%
Symptoms unlikely to be ischemic	1.43%	Thrombolytics	2.32%	CHF complication	2.22%
Unstable angina	45.55%	STEMI	19.72%	CVA/Stroke complication	0.37%
Non-STEMI	22.78%	PCI status: Elective	44.14%	Hemorrhagic stroke complication	0.04%
STEMI	19.72%	PCI status: Emergent	19.51%	Tamponade complication	0.22%
Stable angina	5.65%	BMI	29.97 ± 8.48	New requirement for dialysis complication	0.30%
Cardiogenic shock	3.55%	SBP (mmHg)	118.27 ± 40.64	Other complications	0.31%
Congestive Heart Failure	14.89%	DBP (mmHg)	65.86 ± 23.65	RBC/whole blood transfusion	3.86%
Diabetes	28.51%	Pulse (bpm)	62.01 ± 26.97	Hgb prior to transfusion g/dL	1.33%
Hypertension	77.24%	Troponin T (ng/dL)	0.42 ± 4.94	Bleeding complication w/in 72 hours	1.73%
Dyslipidemia	78.58%	Creatinine (mg/dL)	1.16 ± 2.67	**Discharge Medications**	
Family history of CAD	21.32%	GFR (ml/min/m2)	74.69 ± 58.26	ASA n (%)	93.12%
Current Smoker	19.63%	Hemoglobin (g/dL)	13.19 ± 7.56	Clopidogrel use at discharge n (%)	88.46%
Prior MI	25.49%	PCI for high risk non-STEMI	28,52%	Ticlopidine use at discharge n (%)	0.17%
Prior PCI	24.39%	Immediate PCI for STEMI	7.61%	Prasugrel use at discharge n (%)	0.59%
Prior CABG	17.05%	Staged PCI	0.21%	Ticagrelor use at discharge n (%)	3.69%
Peripheral artery disease	13.88%	PCI for STEMI (Unstable, >12 hours from symptom onset)	0.77%	Statin use at discharge n (%)	86.43%
Cerebrovascular disease	11.15%	PCI for STEMI (stable after successful full-dose thrombolysis))	0.30%		
Chronic Kidney Disease: Severe	5.61%	PCI for STEMI (stable, > 12 from hours symptom onset)	0.27%		
		Rescue PCI for STEMI (after failed full-dose lytics)	0.56%		

Fig. 1: List of clinical variables and rate of positive occurance as %.

2.1 ECG Data Processing

The ECG signals are stored in a standard image (grid paper) format, where a conventional 12-lead ECG is presented in six rows (see main manuscript. Each of the 12 ECG leads signifies a distinct direction of cardiac activation in three-dimensional space. The conventional ECG leads include the limb leads comprised of Leads I, II, III, aVR, aVL, and aVF, and the precordial leads consisting of Leads V1, V2, V3, V4, V5, and V6. The upper three rows demonstrate individual lead signals (I, II, III, aVR, aVF, aVL, and V1-V6), each covering a duration of 2.5 seconds. The lower three rows present a comprehensive 'rhythm strip' featuring various leads, capturing the entirety of the 10-second ECG duration.

The ECG signals corresponding to each lead are cropped from the first three rows of ECG image using an automated Python script (as indicated in Figure **??**(b)). These cropped images are resized to 124x124x3 by keeping the aspect ratio same, to ensure the same input image size for the deep learning models.

3 Explainability and Interpretability of Models

While studying the importance of different EHR features in the prediction of post-PCI prognosis, it was observed that the prior stroke variable was given significant importance by the baseline (EHR-only) model for stroke prediction (as shown in Figure 2.a. Similarly, as indicated in Figure 2.b, the prior status of congestive heart failure is given significant importance in the determination of heart failure post-PCI.

The EHR model indicates the likelihood of a patient experiencing a stroke or heart failure within six months after PCI if there was a prior occurrence of stroke or heart failure. However, it is important to note that this condition alone may not be deemed necessary and sufficient for determining the respective events post-PCI since missing or incomplete history may also cause the model to fail.

To understand the importance of the ECG snippets, the attention rollout method is employed to assess the information flow across the multi-head attention layers within the transformer blocks of the ViT model. This technique quantifies the distribution of attention and aids in understanding how information is processed and propagated throughout the model's attention mechanisms. Figure 2.c represents the attention rollout obtained for true positive and false positive samples of each targeted outcome.

As revealed by the attention distribution results depicted in Figure 2.c, when predicting heart failure, the ViT model notably highlights the significance of the V1 lead in instances of true positive samples. In the context of predicting all-cause mortality, the model's attention varies between leads and pixels in the vicinity of the signals, for different samples. Notably, the V5 lead emerges as of substantial importance when forecasting stroke outcomes.

The CNN+CBAM model is designed as a branched architecture that parses the two input leads separately and afterwards, we concatenate the feature map from two ECG branches. Thus, it is not feasible to identify the attention maps from the CNN+CBAM model.

References

[1] Pablo Juan-Salvadores, Cesar Veiga, Víctor Alfonso Jiménez Díaz, Alba Guitián González, Cristina Iglesia Carreño, Cristina Martínez Reglero, José Antonio Baz Alonso, Francisco Caamaño Isorna, and Andrés Iñiguez Romo. Using

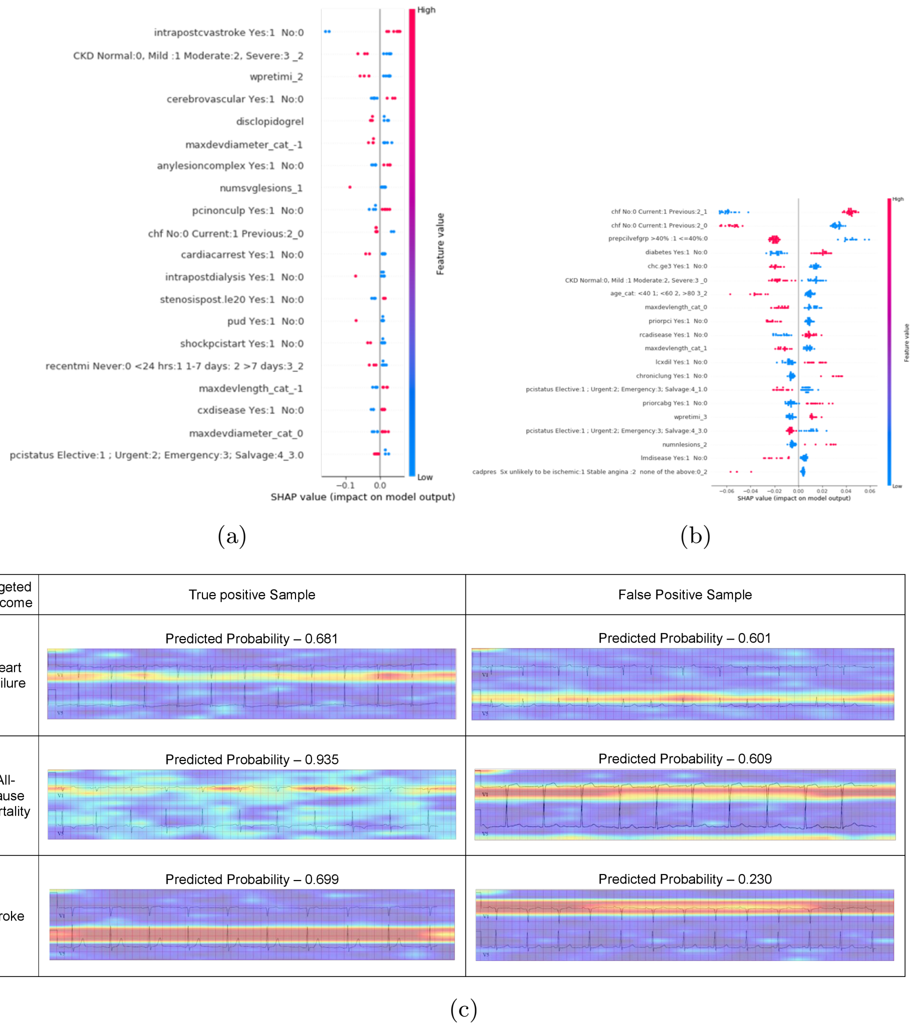

Fig. 2: Model interpretation. Shapely addictive explanations (SHAP) of EHR-only model predictions for (a) stroke, (b) heart failure. (c) Heatmaps - Understanding ECG image reading by the model using attention rollout for various clinical event.

machine learning techniques to predict mace in very young acute coronary syndrome patients. *Diagnostics*, 12(2):422, 2022.

[2] Yijiang Zhou, Ruoyu Zhu, Xiaojun Chen, Xiaolei Xu, Qiwen Wang, Liujun Jiang, Jianhua Zhu, Jian Wu, Hui Yan, and Li Zhang. Machine learning-based cardiovascular event prediction for percutaneous coronary intervention. *Journal of the American College of Cardiology*, 73(9S1):127–127, 2019.

[3] Jinwan Wang, Shuai Wang, Mark Xuefang Zhu, Tao Yang, Qingfeng Yin, and Ya Hou. Risk prediction of major adverse cardiovascular events occurrence within

6 months after coronary revascularization: machine learning study. *JMIR Medical Informatics*, 10(4):e33395, 2022.

[4] Juan-Jose Beunza, Enrique Puertas, Ester García-Ovejero, Gema Villalba, Emilia Condes, Gergana Koleva, Cristian Hurtado, and Manuel F Landecho. Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease). *Journal of biomedical informatics*, 97:103257, 2019.

[5] Micah Liam Arthur Heldeweg, Nan Liu, Zhi Xiong Koh, Stephanie Fook-Chong, Weng Kit Lye, Mark Harms, and Marcus Eng Hock Ong. A novel cardiovascular risk stratification model incorporating ecg and heart rate variability for patients presenting to the emergency department with chest pain. *Critical Care*, 20:1–9, 2016.

[6] Jiandong Zhou, Lakshmi Murugappan, Lei Lu, Oscar Hou In Chou, Bernard MY Cheung, Gary Tse, and Tingting Zhu. Multi-modality machine learning models to predict stroke and atrial fibrillation in patients with heart failure. *medRxiv*, pages 2023–11, 2023.