FEASIBILITY OF INTEGRATING RULE-BASED MODELS WITH MACHINE LEARNING ALGORITHMS FOR THE DETECTION OF FALSE ALARMS ASSOCIATED WITH CARDIAC ARRHYTHMIAS IN INTENSIVE CARE UNITS (ICUS).

HAROLD HERNANDO RODRIGUEZ RODRIGUEZ

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS
ESCUELA DE INGENIERÍA ELÉCTRICA, ELECTRÓNICA Y DE TELECOMUNICACIONES
MAESTRÍA EN INGENIERÍA DE TELECONUMICACIONES
BUCARAMANGA
2026

FEASIBILITY OF INTEGRATING RULE-BASED MODELS WITH MACHINE LEARNING ALGORITHMS FOR THE DETECTION OF FALSE ALARMS ASSOCIATED WITH CARDIAC ARRHYTHMIAS IN INTENSIVE CARE UNITS (ICUS).

HAROLD HERNANDO RODRIGUEZ RODRIGUEZ

Thesis submitted in partial fulfillment of the requirements for the degree of Master in Telecommunications Engineering.

Advisor
Carlos A. Fajardo, Ph.D.

Co-advisor
Hans García, Ph.D

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS
ESCUELA DE INGENIERÍA ELÉCTRICA, ELECTRÓNICA Y DE TELECOMUNICACIONES
MAESTRÍA EN INGENIERÍA DE TELECONUMICACIONES
BUCARAMANGA
2026

CONTENT

LIST OF FIGURES

LIST OF TABLES

ANNEXES

RESUMEN

TÍTULO VIABILIDAD DE INTEGRAR MODELOS BASADOS EN REGLAS CON ALGORITMOS DE APRENDIZAJE AUTOMÁTICO PARA LA DETECCIÓN DE FALSAS ALARMAS ASOCIADAS A ARRITMIAS CARDÍACAS EN UNIDADES DE CUIDADOS INTENSIVOS (UCI). *

AUTOR: HAROLD HERNANDO RODRIGUEZ RODRIGUEZ **

PALABRAS CLAVE: Falsas Alarmas, ICU, Detección, Modelos Basados en Reglas, Modelos de Aprendizaje Automático.

DESCRIPCIÓN: Las falsas alarmas por arritmias en Unidades de Cuidados Intensivos (UCI) pueden alcanzar hasta el 88.8 %, generando fatiga por alarmas y retrasos en la atención. Este trabajo presenta una red siamesa multimodal liviana que integra señales fisiológicas crudas, metadatos de arritmias y salidas de un modelo basado en reglas. Utilizando el conjunto de datos PhysioNet/CinC 2015, el sistema aplica codificadores convolucionales multiescala, una función de pérdida adaptada al desbalance de clases y estrategias de preprocesamiento y aumento de datos.

La validación cruzada de cinco pliegues confirma la relevancia del aumento de datos, los embeddings de arritmias y el conocimiento basado en reglas para mejorar la capacidad discriminativa. El modelo alcanza un **puntaje Challenge de 0.8498** $\pm$ **0.0686** y un **F1-score de 0.9101** $\pm$ **0.0406**, superando los modelos base con casi tres veces menos parámetros (0.8789M vs. 3.28M).

Este estudio demuestra que integrar sistemas basados en reglas con modelos de deep learning es viable y beneficioso, al incorporar conocimiento médico mediante capas embedding y funciones de pérdida especializadas. Los resultados evidencian un marco eficiente para reducir falsas alarmas en tiempo real en UCI, resaltando el potencial del aprendizaje multimodal y la optimización basada en conocimiento clínico en IA médica.

---

* Tesis

** Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingeniería Eléctrica, Electrónica y de Telecomunucaciones. Director: Carlos A. Fajardo, Ph.D. Codirector: Hans García, Ph.D.

ABSTRACT

TITLE: FEASIBILITY OF INTEGRATING RULE-BASED MODELS WITH MACHINE LEARNING ALGORITHMS FOR THE DETECTION OF FALSE ALARMS ASSOCIATED WITH CARDIAC ARRHYTHMIAS IN INTENSIVE CARE UNITS (ICUS). *

AUTOR: HAROLD HERNANDO RODRIGUEZ RODRIGUEZ **

Keywords: False Alarms, ICU, Detection, Rule based models, Deep learning models.

Description: False arrhythmia alarms in Intensive Care Units (ICUs) can reach rates as high as 88.8 %, leading to alarm fatigue and delayed clinical response. This work presents a lightweight multimodal Siamese network that integrates raw physiological signals, arrhythmia metadata, and outputs from a rule-based model. Using the PhysioNet/CinC 2015 dataset, the system employs multi-scale convolutional encoders, a loss function adapted to class imbalance, and tailored preprocessing and data augmentation strategies.

Five-fold cross-validation confirms the relevance of data augmentation, arrhythmia embeddings, and rule-based knowledge in improving discriminative performance. The proposed model achieves a **Challenge score of 0.8498** $\pm$ **0.0686** and an **F1-score of 0.9101** $\pm$ **0.0406**, outperforming baseline models with nearly three times fewer parameters (0.8789M vs. 3.28M).

This study demonstrates that integrating rule-based systems with deep learning models is both feasible and beneficial, as it incorporates medical knowledge through embedding layers and specialized loss functions. The results highlight an efficient framework for reducing false alarms in real time in the ICU, underscoring the potential of multimodal learning and clinical knowledge-based optimization in medical AI.

---

* Thesis

** Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingeniería Eléctrica, Electrónica y de Telecomunucaciones. Advisor: Carlos A. Fajardo, Ph.D. Electronic Engineer. Co-advisor: Hans García, Ph.D.

INTRODUCTION

Intensive Care Units (ICUs) are organized systems designed to provide care for critically ill patients, offering intensive and specialized medical and nursing care. These ICUs have enhanced monitoring capabilities and various modalities of physiological organ support to sustain life during periods of life-threatening organ failure [1]. The monitoring of vital signs is conducted using bedside monitors, which sense vital signs and various patient-related variables. Among the monitored signals are variables such as the pulse oximeter (PPG), electrocardiogram (ECG), arterial blood pressure catheter (ABP), and central venous pressure catheter, among others. These monitors are equipped with an alarm system that alerts the staff when physiological signals are outside the predefined ranges. Specifically, arrhythmia alarms on these monitors are designed to be highly sensitive and not miss potentially life-threatening events. The problem with this high sensitivity is that it compromises the specificity of the alarms [2], resulting in a false alarm rate that can reach up to 88.8 % [3].

Alarms can be activated by numerous factors, such us patient movement, power line interference, electrode contact noise, and noise from data collection devices. Falsely triggered alarms pose a hidden danger in ICUs, as they not only lead to sleep deprivation but also put patients at risk of desensitization to warnings and slowing

---

[1] John C. Marshall et al. «What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine». En: *Journal of critical care* 37 (feb. de 2017), págs. 270-276. DOI: 10.1016/J.JCRC.2016.07.015.

[2] Barbara J. Drew et al. «Practice Standards for Electrocardiographic Monitoring in Hospital Settings». En: *Circulation* 110 (17 oct. de 2004), págs. 2721-2746. DOI: 10.1161/01.CIR.0000145144.56673.59.

[3] Barbara J. Drew et al. «Insights into the Problem of Alarm Fatigue with Physiologic Monitor Devices: A Comprehensive Observational Study of Consecutive Intensive Care Unit Patients». En: *PLOS ONE* 9 (10 oct. de 2014), e110274. DOI: 10.1371/JOURNAL.PONE.0110274.

response times [4] [5]. On the other hand, only between 2 % and 9 % of all ICU alarms are correctly triggered, and only this percentage truly requires a response from the medical staff [6].

Rule-based techniques employ predefined sets of logical conditions to assess alarm validity, providing simplicity and interpretability [7] [8] [9]. However, these techniques have limitations, such as the inability to adapt to unexpected variations and a reliance on the quality of input signals, which can result in a high false alarm rate or the omission of critical events [10] [11]. In contrast, machine learning and deep learning approaches have shown significant improvements in false alarm detection. Algorithms such

---

[4] Sairam Parthasarathy y Martin J. Tobin. «Sleep in the intensive care unit». En: *Intensive Care Medicine* 30 (2 feb. de 2004), págs. 197-206. DOI: 10.1007/S00134-003-2030-6/FIGURES/4.

[5] Wynne E. Morrison et al. *Noise, stress, and annoyance in a pediatric intensive care units : Critical Care Medicine*. 2003.

[6] Christine L Tsien y James C Fackler. *Poor prognosis for existing monitors in the intensive care units : Critical Care Medicine*. 1997.

[7] Charalampos Tsimenidis y Alan Murray. «Reliability of clinical alarm detection in intensive care units». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 1185-1188. DOI: 10.1109/CIC.2015.7411128.

[8] Sibylle Fallet, Sasan Yazdani y Jean Marc Vesin. «A multimodal approach to reduce false arrhythmia alarms in the intensive care unit». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 277-280. DOI: 10.1109/CIC.2015.7408640.

[9] Sibylle Fallet, Sasan Yazdani y Jean Marc Vesin. «False arrhythmia alarms reduction in the intensive care unit: a multimodal approach». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1217. DOI: 10.1088/0967-3334/37/8/1217.

[10] Runnan He et al. «Reducing false arrhythmia alarms in the ICU using novel signal quality indices assessment method». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 1189-1192. DOI: 10.1109/CIC.2015.7411129.

[11] Erdem Yanar y Yesim Serinagaoglu Dogrusoz. «False Ventricular-Fibrillation/Flutter Alarm Reduction of Patient Monitoring Systems in Intensive Care Units». En: *MeMeA 2018 - 2018 IEEE International Symposium on Medical Measurements and Applications, Proceedings* (ago. de 2018). DOI: 10.1109/MEMEA.2018.8438601.

as Naive Bayes, SVM [12], decision trees [13] [14], and convolutional neural networks [15] [16], have demonstrated high positive classification rates and precision in arrhythmia detection. Despite, these improvements still encounter challenges related to data quality and availability, generalization, and interpretability.

The integration of rule-based models with machine learning techniques has emerged as a promising strategy, combining the robustness of rules with the adaptability and precision of machine learning algorithms. This hybrid approach can better handle the variability of physiological signals and improve the accuracy of arrhythmia detection [17] [18], although challenges remain in optimizing and interpreting the resulting

---

[12] Benedikt Baumgartner, Kolja Rodel y Alois Knoll. «A data mining approach to reduce the false alarm rate of patient monitors». En: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2012), págs. 5935-5938. DOI: 10.1109/EMBC.2012.6347345.

[13] Miguel Caballero y Grace M. Mirsky. «Reduction of false cardiac arrhythmia alarms through the use of machine learning techniques». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 1169-1172. DOI: 10.1109/CIC.2015.7411124.

[14] Tishya Manna, Aleena Swetapadma y Moloud Abdar. «Decision Tree Predictive Learner-Based Approach for False Alarm Detection in ICU». En: *Journal of Medical Systems* 43 (7 jul. de 2019), págs. 1-13. DOI: 10.1007/S10916-019-1337-Y/TABLES/5.

[15] Martin Zihlmann, Dmytro Perekrestenko y Michael Tschannen. «Convolutional recurrent neural networks for electrocardiogram classification». En: *Computing in Cardiology* 44 (2017), págs. 1-4. DOI: 10.22489/CINC.2017.070-060.

[16] Qiang Yu et al. «Intensive Care Unit False Alarm Identification Based on Convolution Neural Network». En: *IEEE Access* 9 (2021), págs. 81841-81854. DOI: 10.1109/ACCESS.2021.3086862.

[17] Sardar Ansari, Ashwin Belle y Kayvan Najarian. «Multi-modal integrated approach towards reducing false arrhythmia alarms during continuous patient monitoring: The Physionet Challenge 2015». En: *Computing in Cardiology* 42 (feb. de 2015). DOI: 10.1109/CIC.2015.7411127.

[18] V. Kalidas y L. S. Tamil. «Cardiac arrhythmia classification using multi-modal signal analysis». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1253. DOI: 10.1088/0967-3334/37/8/1253.

decisions [19].

This research aims to determine the feasibility of integrating rule-based models with learning algorithms to detect false alarms associated with cardiac arrhythmias in intensive care units. To this end, an evaluation of the integrated model's performance will be conducted.

[19] Shenda Hong et al. «ENCASE: An ENsemble ClASsifiEr for ECG classification using expert features and deep neural networks». En: *Computing in Cardiology* 44 (2017), págs. 1-4. DOI: 10.22489/CINC.2017.178-245.

State of the Art

Early efforts to reduce false alarms in ICUs focused on enforcing physiological coherence across signals and explicitly assessing signal quality. Pioneering studies proposed the joint use of invasive arterial blood pressure (ABP) and photoplethysmography (PPG) to mechanically corroborate ECG-based alarms, suppressing those lacking hemodynamic evidence and thereby reducing false positives caused by artifacts or lead disconnections [20] [21]. In parallel, data-fusion frameworks and signal quality indices (SQIs) were introduced as safety mechanisms to determine when each channel could be considered reliable [22]. The analysis of ECG quality during arrhythmias further demonstrated the value of SQIs as abstention criteria under noisy conditions [23]. Additionally, clinical demonstrations of inter-sensor correlation logic in real ICU environments reinforced the operational feasibility of these physiologically grounded strategies [24].

---

[20] Gari D Clifford et al. «Using the blood pressure waveform to reduce critical false ECG alarms | IEEE Conference Publication | IEEE Xplore». En: *2006 Computers in Cardiology*. IEEE, 2006, págs. 829-832.

[21] Anton Aboukhalil et al. «Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform». En: *Journal of Biomedical Informatics* 41 (3 jun. de 2008), págs. 442-451. DOI: 10.1016/J.JBI.2008.03.003.

[22] Qiao Li y Gari D. Clifford. «Signal quality and data fusion for false alarm reduction in the intensive care unit». En: *Journal of Electrocardiology* 45 (6 nov. de 2012), págs. 596-603. DOI: 10.1016/J.JELECTROCARD.2012.07.015.

[23] Joachim Behar et al. «ECG signal quality during arrhythmia and its application to false alarm reduction». En: *IEEE Transactions on Biomedical Engineering* 60 (6 2013), págs. 1660-1666. DOI: 10.1109/TBME.2013.2240452.

[24] Yuval Bitan y Michael F. O'Connor. «Correlating data from different sensors to increase the positive predictive value of alarms: an empiric assessment». En: *F1000 Research 2012 1:45* 1 (nov. de 2012), pág. 45. DOI: 10.12688/f1000research.1-45.v1.

A major milestone in the context of false arrhythmia alarm reduction was the PhysioNet/Computing in Cardiology Challenge 2015, which formalized the problem and accelerated progress in the field. This competition standardized the alarm categories into five clinically relevant types—asystole, extreme bradycardia, extreme tachycardia, ventricular fibrillation/flutter, and ventricular tachycardia—and divided the task into two scenarios ("real-time" and "retrospective"), while introducing a scoring metric that strongly penalized missed true alarms [25]. By establishing a common benchmark, the challenge enabled fair comparisons across methods and fostered rapid methodological advances. As a result, a wide range of solutions emerged, including rule-based systems, machine learning approaches, and multimodal fusion strategies that leveraged complementary information from ECG, ABP, and PPG channels [8] [26] [27].

Following the Physionet challenge, a new wave of deterministic and rule-based algorithms emerged, focusing on low-latency suppression of false alarms through explicit physiological criteria and SQI-driven decisions. These approaches commonly verified critical ECG alarms through mechanical evidence in ABP and PPG signals to mitigate artefact-related false detections, and proposed practical multimodal rules that combined ECG with ABP and/or PPG [28]. In parallel, several works introduced

---

[25] Gari D. Clifford et al. «The PhysioNet/Computing in Cardiology Challenge 2015: Reducing false arrhythmia alarms in the ICU». En: *Computing in Cardiology* 42 (2015). Cited by: 116; All Open Access, Green Open Access, 273 – 276. DOI: 10.1109/CIC.2015.7408639.

[26] Filip Plesinger et al. «False alarms in intensive care unit monitors: Detection of life-threatening arrhythmias using elementary algebra, descriptive statistics and fuzzy logic». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 281-284. DOI: 10.1109/CIC.2015.7408641.

[27] Chengyu Liu, Lina Zhao y Hong Tang. «Reduction of False Alarms in Intensive Care Unit using Multi-feature Fusion Method». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 741-744. DOI: 10.1109/CIC.2015.7411017.

[28] Wei Zong. «Reduction of false critical ECG alarms using waveform features of arterial blood pressure and/or photoplethysmogram signals». En: *Computing in Cardiology* 42 (feb. de 2015),

real-time frameworks that integrated ECG quality assessment with pulse monitoring [29], as well as fuzzy-logic systems and transparent morphological rule sets designed for life-threatening arrhythmias, emphasizing interpretability and operational safety [26] [30]. Additionally, robust QRS detectors were developed for noisy ICU environments [31]. Across these efforts, a defining characteristic was the systematic use of modality-specific SQIs as a safety layer and abstention policy, ensuring that alarm decisions were made only when the available signals were deemed reliable [9] [10].

Beyond rule-driven approaches, numerous studies addressed the problem from a purely data-driven perspective, relying on engineered features or learned representations. From a classical machine learning standpoint, multimodal feature sets—capturing waveform morphology, heart-rate variability, ECG-pulse coherence, and ABP integrity—were combined with classifiers such as Random Forests, Support Vector Machines, or boosted trees [31] [32] [33]. Notably, an optimized Random Forest incorporating feature selection and SQI analysis achieved one of the top scores in the real-time

---

págs. 289-292. DOI: 10.1109/CIC.2015.7408643.

[29] Vessela Krasteva et al. «Real-time arrhythmia detection with supplementary ECG quality and pulse wave monitoring for the reduction of false alarms in ICUs». En: *Physiological measurement* 37 (8 jul. de 2016), págs. 1273-1297. DOI: 10.1088/0967-3334/37/8/1273.

[30] F. Plesinger et al. «Taming of the monitors: reducing false alarms in intensive care units». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1313. DOI: 10.1088/0967-3334/37/8/1313.

[31] Nadi Sadr et al. «Reducing false arrhythmia alarms in the ICU by Hilbert QRS detection». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 1173-1176. DOI: 10.1109/CIC.2015.7411125.

[32] Sardar Ansari et al. «Suppression of false arrhythmia alarms in the ICU: a machine learning approach». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1186. DOI: 10.1088/0967-3334/37/8/1186.

[33] Qiang Zhang et al. «Reducing false arrhythmia alarm rates using robust heart rate estimation and cost-sensitive support vector machines». En: *Physiological Measurement* 38 (2 ene. de 2017), pág. 259. DOI: 10.1088/1361-6579/38/2/259.

category of the Challenge (83.08) on the hidden test set [34]. Promising results were also obtained using ABP signal alone, including suppression of 90.3 % of false extreme tachycardia alarms with only 0.54 % of true alarms incorrectly rejected [35], as well as improved performance through class-probability assignment strategies in Random Forest classifiers [36].

With the advent of deep learning, end-to-end architectures began to replace manual feature design. CNN–LSTM models with attention mechanisms fusing ECG, ABP, and PPG demonstrated high sensitivity and specificity [37], while grouped-convolution ensembles obtained strong scores on the Challenge hidden test set [16]. Additional efforts explored early-classification pipelines based on ResNet and BiLSTM architectures [38] [39] [40]. More recently, generative diffusion models conditioned on pre-alarm

---

[34] Wan Tai M. Au-Yeung et al. «Reduction of false alarms in the intensive care unit using an optimized machine learning based approach». En: *npj Digital Medicine* 2 (1 dic. de 2019), págs. 1-5. DOI: 10.1038/S41746-019-0160-7;SUBJMETA.

[35] Petre Lameski et al. «Suppression of Intensive Care Unit False Alarms Based on the Arterial Blood Pressure Signal». En: *IEEE Access* 5 (2017), págs. 5829-5836. DOI: 10.1109/ACCESS.2017.2690380.

[36] Krzysztof Gajowniczek, Iga Grzegorczyk y Tomasz Ząbkowski. «Reducing False Arrhythmia Alarms Using Different Methods of Probability and Class Assignment in Random Forest Learning Methods». En: *Sensors 2019, Vol. 19, Page 1588* 19 (7 abr. de 2019), pág. 1588. DOI: 10.3390/S19071588.

[37] Sajad Mousavi, Atiyeh Fotoohinasab y Fatemeh Afghah. «Single-modal and multi-modal false arrhythmia alarm reduction using attention-based convolutional and recurrent neural networks». En: *PLOS ONE* 15 (1 ene. de 2020), e0226990. DOI: 10.1371/JOURNAL.PONE.0226990.

[38] Eric P. Lehman et al. «Representation Learning Approaches to Detect False Arrhythmia Alarms from ECG Dynamics». En: *Proceedings of machine learning research* 85 (2018), pág. 571.

[39] Jack Boynton y Byung Suk Lee. «Deep Learning Based Classification of True/False Arrhythmia Alarms in the Intensive Care Unit». En: *Computing in Cardiology* 2021-September (2021). DOI: 10.23919/CINC53138.2021.9662874.

[40] Karol Ardila et al. «Detecting False Arrhythmias Alarms in the ICU Using a Deep Learning Approach». En: *2024 XXIV Symposium of Image, Signal Processing, and Artificial Vision (STSIVA)*. 2024, págs. 1-5. DOI: 10.1109/STSIVA63281.2024.10637887.

signal context have shown potential by discriminating true and false alarms through prediction–observation discrepancy [41]

A parallel line of research has focused on hybrid strategies that combine physiologically inspired rules and SQI analysis with learned models to enhance robustness in alarm classification. Within this category, systems that fuse handcrafted rule sets with CNN-based architectures have been shown to outperform pure CNN models, achieving strong event-level performance while preserving multimodal flexibility [42]. Complementary efforts have incorporated rule-derived embeddings into CNN pipelines, together with contrastive objectives, to improve the separability between true and false alarms [43]. Likewise, ML frameworks that integrate SQIs and abstention logic have demonstrated that embedding clinical physiological knowledge can maintain high sensitivity while reducing alarm burden [34] [9]. Recent extensions toward multimodal Siamese architectures further emphasize similarity learning and cross-channel coherence, yielding performance gains over comparable baselines [44]. Collectively, these studies suggest that hybrid approaches—combining rules with ML/DL—provide the most effective compromise between false alarm reduction, clinical sensitivity, and real-time operability [43] [42] [44]

---

[41] Feng Wu et al. *A Diffusion Model with Contrastive Learning for ICU False Arrhythmia Alarm Reduction*. 2023.

[42] Sandeep Chandra Bollepalli et al. «Real-Time Arrhythmia Detection Using Hybrid Convolutional Neural Networks». En: *Journal of the American Heart Association* 10 (23 dic. de 2021), pág. 23222. DOI: 10.1161/JAHA.121.023222.

[43] Yuerong Zhou et al. «A contrastive learning approach for ICU false arrhythmia alarm reduction». En: *Scientific Reports 2022 12:1* 12 (1 mar. de 2022), págs. 1-10. DOI: 10.1038/s41598-022-07761-9.

[44] Harold H. Rodriguez, Hans Y. Garcia y Carlos A. Fajardo. «Embedding-Enhanced Multimodal Siamese Network for False Arrhythmia Alarm Reduction in Intensive Care Units». En: *2025 25th Symposium of Image, Signal Processing, and Artificial Vision, STSIVA 2025* (2025). DOI: 10.1109/STSIVA66383.2025.11156764.

Among the most relevant and recent research in false alarm prediction in ICUs, an innovative approach stand out. In 2022, Zhou et al. [43] proposed a deep learning framework that employs convolutional neural networks (CNNs) to automatically learn feature representations from physiological waveforms. Their method combines contrastive learning to enhance discrimination between true and false alarms with learned embeddings derived from a rule-based system, integrating prior clinical domain knowledge for each alarm type, and achieved a Challenge score of 84.47.

Although this study builds upon the foundations established by Zhou et al. [43], our approach introduces several key contributions. Specifically, we implement embedding layers to represent arrhythmia-type information and rule-based outputs more efficiently, propose a novel distance-based loss function instead of an angular one, and design a more compact architecture with fewer trainable parameters. Additionally, we incorporate a tailored data augmentation strategy that proves particularly effective given the nature of the proposed Siamese-based architecture.

1. OBJECTIVES

1.1. GENERAL OBJECTIVE

To evaluate the feasibility of integrating rule-based models with machine learning algorithms in order to improve the detection of false alarms in cardiac arrhythmia monitoring systems in intensive care units (ICUs) settings.

1.2. SPECIFIC OBJECTIVES

- To preprocess and structure the database to facilitate the training of models for detecting false alarms associated with cardiac arrhythmias.
- To select a suitable rule-based model and a machine learning model for detecting false alarms related to cardiac arrhythmias in Intensive Care Units (ICUs).
- To develop an integrated model that combines the rule-based approach with a machine learning algorithm to improve the performance in the detection of false alarms associated with cardiac arrhythmias.
- To evaluate the performance of the integration of both models to determine the feasibility of their application in the detection of false alarms associated with cardiac arrhythmias in Intensive Care Units (ICUs).

2. METHODS

2.1. FRAMEWORK OVERVIEW

Figure 2 illustrates the overall structure of the proposed framework. It consists of three main components: (i) a dual-branch Siamese encoder that extracts discriminative representations from paired signal segments, (ii) a multimodal embedding module that integrates arrhythmia-type and rule-based information, and (iii) a fusion and classification stage that estimates the probability of an alarm being true or false. This modular design allows the model to jointly exploit temporal, morphological, and contextual features while maintaining a lightweight computational footprint suitable for real-time monitoring.

The proposed multimodal Siamese framework is designed to detect false arrhythmia alarms in intensive care units by integrating heterogeneous sources of information within a unified deep learning architecture. The system processes multi-channel physiological signals in parallel with categorical metadata and the output of a rule-based model, combining data-driven learning with prior domain knowledge to improve alarm reliability.

Unlike task-specific deep networks, the proposed framework is formulated to operate on any multimodal physiological dataset that provides synchronized signal channels and associated categorical metadata. Its Siamese configuration enforces a consistent latent representation between corresponding signal pairs, encouraging the model to learn discriminative patterns that generalize across patients, acquisition devices, and monitoring conditions.

Overall, this framework establishes a general methodology for multimodal alarm validation, where raw signals, categorical inputs, and expert-derived information are combined in a scalable, interpretable, and dataset-agnostic manner.

2.2. DATA PREPROCESSING PIPELINE

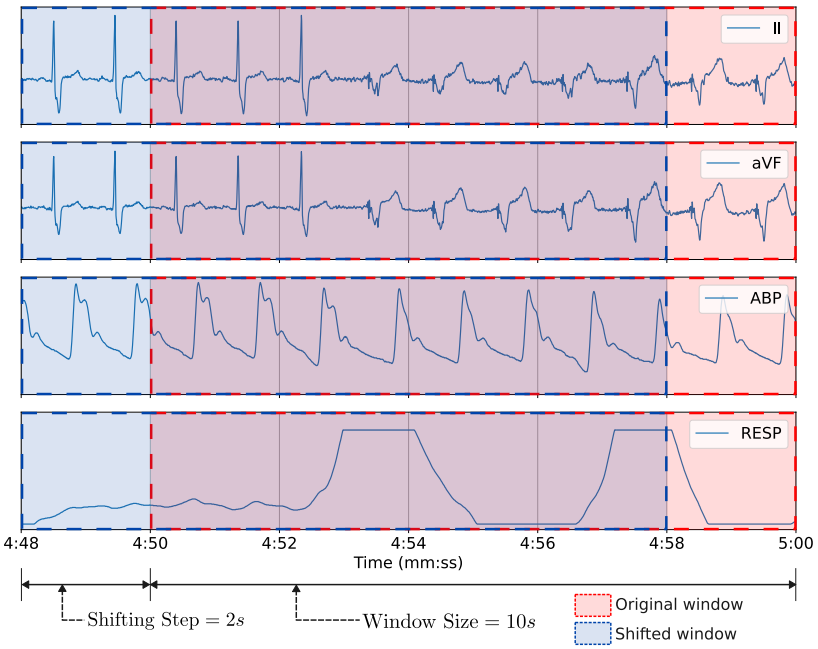
The data preprocessing pipeline was designed to ensure uniformity, temporal coherence, and compatibility of multimodal physiological signals for deep neural processing. Each record was first standardized in terms of signal duration and channel count to guarantee consistent tensor dimensions across samples. When a particular physiological signal was missing, a placeholder zero-filled waveform was inserted to preserve structural alignment among channels. All signals were subsequently normalized using the z-score transformation to enforce zero mean and unit variance, mitigating inter-channel amplitude disparities and improving training stability.

Within the proposed framework, each input sample consists of two temporally related signal segments: the *alarm* segment, representing the period associated with the triggered event, and the *reference* segment, extracted from a different temporal context of the same recording. This pairwise formulation enables the Siamese network to learn relational representations that emphasize similarity for false alarms and dissimilarity for true alarms. Importantly, this pairing process does not constitute data augmentation; rather, it is an inherent learning strategy that structures the latent space according to physiological correspondence.

Figure 1 shows how a temporal data augmentation strategy based on window shifting was applied to explicitly enhance robustness and variability. In this approach, the original analysis window (red) is complemented by an additional window displaced by a small temporal offset (blue), increasing tolerance to temporal misalignment and noise. For the reference segments, random sampling is performed within valid temporal regions that do not overlap with the alarm window, ensuring independent contextual information for each pair. Categorical and rule-based inputs are consistently duplicated across augmented instances to maintain alignment among modalities. Overall, this integrated procedure provides a general and reproducible strategy for preparing multimodal physiological recordings, strengthening the robustness and generaliza-

tion capabilities of the model across diverse ICU monitoring environments.

Figure 1. Illustration of the temporal window-shifting strategy



Note: The original analysis window (red) is complemented by a shifted version (blue) displaced by a small temporal offset. This augmentation technique increases variability and enhances robustness to temporal misalignment and noise.

2.3. RULE-BASED MODEL

In this study, we used the rule-based model proposed by Plesinger et al. [26], which was designed to minimize false alarms in intensive care units. This approach uses elementary algebra, descriptive statistics, and fuzzy logic to analyze five-minute multimodal records consisting of two ECG channels and one or two blood pressure or plethysmography channels. The model primarily examines the final 14 seconds of the signal because arrhythmia is most likely to occur during this time. First, invalid segments of the records are identified and discarded. Then, QRS complexes are detected, and cardiac activity is evaluated for regularity. If any of the channels show a regular rhythm, the alarm is considered false. Otherwise, specific tests are applied

to each type of arrhythmia.

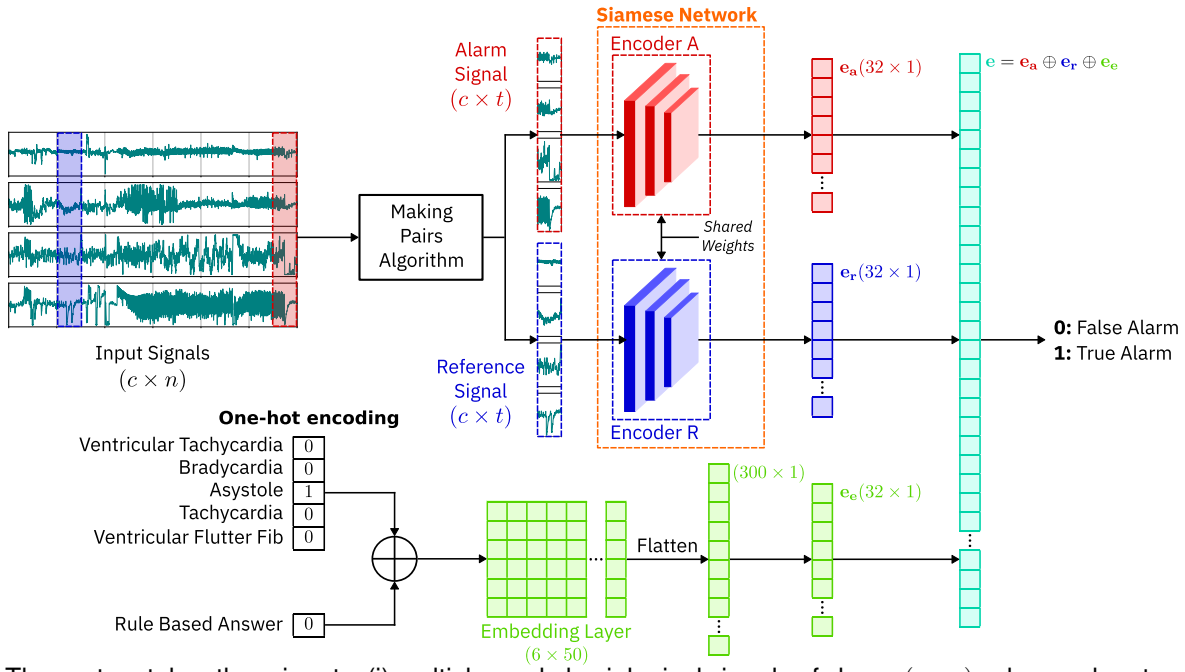
The model includes specific rules to detect asystole, bradycardia, tachycardia, ventricular tachycardia, and ventricular fibrillation/flutter. These rules are based on the temporal distribution of QRS complexes, comparison of heart rate with predefined thresholds, and identification of low-frequency activity in the electrocardiogram (ECG). The model achieved high sensitivities and specificities in the training of 96 % and 89 %, respectively. In the context of our study, we use this model as a reference with which to contrast and complement the results obtained using the proposed multimodal neural network approach.

2.4. MULTIMODAL SIAMESE ARCHITECTURE

Figure 2 illustrates the proposed architecture, which is inspired by the model introduced in [43]. The proposed model receives three different types of input. The first input corresponds to the raw physiological waveforms provided by bedside monitors (*Input Signals*). The second input (*Arrhythmia type vector one-hot encoding*) corresponds to the arrhythmia type associated with the triggered alarm, also provided by the monitor. Finally, the third input corresponds to the output of the rule-based model, which serves as an additional source of prior knowledge for the final decision.

The raw input signals (*Input Signals*), represented by a tensor of shape $(c \times n)$, where c denotes the number of input channels (e.g., different physiological modalities) and n the total number of temporal samples per channel, are first processed by a pair-extraction algorithm, which takes each record and generates two corresponding segments. The first element of the pair is the *alarm signal*, defined as a fixed-length temporal window associated with the alarm event, with shape $(c \times t)$, where t indicates the number of samples within each analysis window. The second element is the *reference signal*, also of shape $(c \times t)$, sampled from a distinct valid temporal region of the same recording. This strategy ensures that both signals represent different

Figure 2. Architecture of the proposed model



The system takes three inputs: (i) multichannel physiological signals of shape $(c \times n)$, where c denotes the number of signal channels and n the number of temporal samples per channel; these are split into alarm segments (red) and reference segments (blue) of shape $(c \times t)$, with t representing the number of samples in each analysis window; (ii) a one-hot encoded arrhythmia type; and (iii) the output of a rule-based model. Alarm and reference signals are processed by a Siamese network with shared encoders, producing feature vectors $\mathbf{e_a}, \mathbf{e_r} \in \mathbb{R}^{128}$. In parallel, the arrhythmia type and rule-based output are embedded into $\mathbf{e_e} \in \mathbb{R}^{128}$. The three vectors are concatenated into $\mathbf{e} \in \mathbb{R}^{384}$ and classified by a sigmoid layer as true or false alarm.

temporal contexts of the same physiological episode.

The alarm and reference signals are then fed into the Siamese network, which is composed of two encoders: Encoder A and Encoder B. These encoders share their weights, ensuring that both inputs are processed under identical conditions. The alarm signal passes through Encoder A and the reference signal through Encoder B, producing the corresponding feature vectors $\mathbf{e_a}$ and $\mathbf{e_r}$.

Both Encoder A and Encoder B, implemented as the `Siamese Network` module. Each encoder is composed of four parallel convolutional blocks, with kernel sizes of 50, 100, 200, and 400, respectively. This multi-scale design enables the model to capture temporal dependencies at different resolutions, ranging from short to long temporal contexts. Each convolutional block consists of two one-dimensional convolutional layers with 32 filters, stride 5, and padding proportional to the kernel size, followed by batch normalization, ReLU activation, dropout with a rate of 0.75, and an adaptive max-pooling operation that reduces the temporal dimension to a single representative value. The outputs of the four convolutional blocks are flattened and concatenated, resulting in a joint representation that is then projected into a 32-dimensional latent vector through a fully connected layer. Finally, a ReLU activation and dropout regularization are applied. The resulting embeddings correspond to the feature vectors for each input signal, denoted as $\mathbf{e_a}$ for the alarm signal and $\mathbf{e_r}$ for the reference signal, with $\mathbf{e_a}, \mathbf{e_r} \in \mathbb{R}^{32}$.

In parallel, the one-hot encoded vector representing the arrhythmia type is concatenated with the output of the rule-based model. The rationale behind this concatenation is to allow the embedding layer to capture potential relationships between the arrhythmia category identified by the bedside monitor and the decision suggested by the rule-based algorithm. The embedding layer is designed with six input units and 50 output dimensions, providing a dense representation of these categorical features. The output of the embedding layer is then flattened and passed through a fully

connected layer with 128 neurons. This final representation is denoted as $\mathbf{e_e}$, where $\mathbf{e_e} \in \mathbb{R}^{32}$.

Finally, the feature vectors $\mathbf{e_a}$, $\mathbf{e_r}$, and $\mathbf{e_e}$ are concatenated into a single joint representation. This concatenated vector is then fed into a classification layer, implemented as a fully connected dense layer with a single neuron and a sigmoid activation function. The output of this layer corresponds to the probability that the alarm is true, enabling the model to perform binary classification between true and false alarms.

2.5. LOSS FUNCTION DESIGN

The proposed model (see Figure 1) incorporates a siamese network into its architecture. The fundamental principle of this type of network is to process two input signals in parallel: the alarm signal $\mathbf{A}$ and the reference signal $\mathbf{R}$, both with dimension $(c \times t)$. Each signal is fed into a convolutional encoder, which ensures that feature extraction is performed under the same representation function. As a result, two latent vectors are obtained: $\mathbf{e_a}$, corresponding to the alarm signal, and $\mathbf{e_r}$, corresponding to the reference signal.

The objective of the Siamese network is to structure the latent space so that the relationship between $\mathbf{e_a}$ and $\mathbf{e_r}$ reflects the nature of the alarm. To achieve this, a composite cost function is used:

$$L = L_{BCE} + wC \tag{1}$$

where L_{BCE} corresponds to binary cross-entropy, which guides the overall training of the model and ensures classification capability. The second term, C, constitutes a Siamese constraint that acts on the relationship between the feature vectors. In particular, if the alarm is false, the network must induce a high similarity between $\mathbf{e_a}$ and $\mathbf{e_r}$; in contrast, if the alarm is true, both vectors must differ significantly in the latent space. The hyperparameter w controls the relative influence of this constraint

on global optimization.

In this research, two variants of the constraint C within the cost function were evaluated: (1) a previously published approach, hereafter referred to as the Zhou Constraint **Zhou Constraint** [43], and (2) our own formulation, referred to as the **Proposed Constraint**.

Zhou Constraint. In this approach, the authors propose a contrastive constraint that operates on the inner product between the latent vectors $\mathbf{e_a}$ and $\mathbf{e_r}$. The central idea is that, in the case of a false alarm, both vectors must have a high degree of similarity (angle close to zero), while in the case of a true alarm, the vectors must differ significantly (angle greater than zero). Mathematically, the constraint is defined as follows:

- For a record i corresponding to a false alarm:

$$C_{false}^{(i)} = -\log\left(\sigma\left(\mathbf{e_a}^T \cdot \mathbf{e_r}\right)\right) \tag{2}$$

  where σ represents the sigmoid function.

- For a record j corresponding to a true alarm:

$$C_{true}^{(j)} = -\log\left(\sigma\left(-\left(\mathbf{e_a}^T \cdot \mathbf{e_r}\right)\right)\right) \tag{3}$$

In this way, the restriction penalizes representations inconsistent with the nature of the alarm, favoring latent similarity in false alarms and dissimilarity in true alarms. Finally, the total contribution of the restriction on a mini-batch is calculated as:

$$C = \frac{1}{N_1}\sum_i C_{false}^{(i)} + \frac{1}{N_2}\sum_j C_{true}^{(j)} \tag{4}$$

where N_1 and N_2 represent the number of false and true alarm records, respectively, in the mini-batch considered.

Proposed Constraint. In this research, we propose a constraint C that combines elements of the contrastive loss traditionally used in Siamese networks with focal loss, in order to improve learning in unbalanced class scenarios. The central idea is that, when the alarm is false, the Euclidean distance between vectors $\mathbf{e_a}$ and $\mathbf{e_r}$ should be reduced; in contrast, when the alarm is true, this distance should be large. The mathematical formulation of this constraint is:

$$C = \frac{1}{N}\sum_{i=1}^{n}(1-y_i)D^2 + \beta\, y_i\, (\mathrm{máx}(0, \alpha - D))^2 \tag{5}$$

where y_i is the sample label ($y_i = 0$ for false alarms and $y_i = 1$ for true alarms), N is the total number of records in the mini-batch, and β is a weighting factor inspired by focal loss, used to give greater emphasis to positive samples (true alarms) due to class imbalance in the dataset. The parameter α defines a minimum margin of separation in the latent space between true and false alarm samples.

The Euclidean distance D between the feature vectors is calculated as:

$$D = \sqrt{\sum_{i=1}^{n}(\mathbf{e_a} - \mathbf{e_r})^2} \tag{6}$$

In this way, the Proposed Constraint not only promotes effective separation between true and false alarms, but also introduces an attention mechanism that compensates for the inherent imbalance in the dataset.

3. EXPERIMENTAL SETUP AND RESULTS

3.1. DATASET DESCRIPTION

This research uses the PhysioNet/Computing in Cardiology Challenge 2015 database: Reducing False Arrhythmia Alarms in the ICU [25]. It is an international reference repository designed to study false alarms in intensive care units (ICUs). The database contains multi-channel records of critically ill patients obtained from four hospitals in the United States and Europe using equipment from three different bedside monitor manufacturers, ensuring clinical and technological diversity in the data.

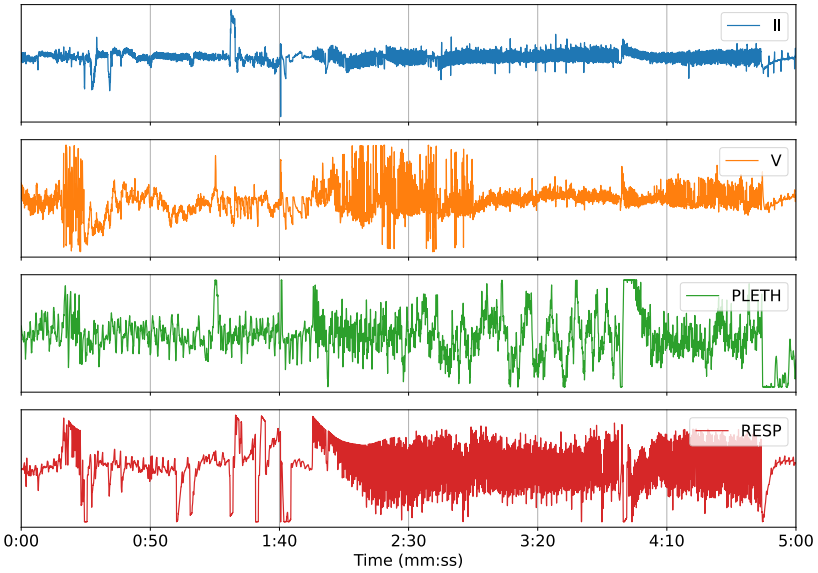
Figure 3 illustrates an example of a five-minute record from the dataset, corresponding to a false asystole alarm. The segment includes four physiological channels: ECG lead II (blue, top panel), ECG lead V (orange), photoplethysmography (green), and respiration (red). Variations in morphology and artefacts typical of the ICU environment are evident, highlighting the complexity of distinguishing between true and false alarms.

Each record is associated with an episode in which the monitor identified a potentially lethal arrhythmia. For each case, at least two electrocardiogram (ECG) leads were stored, as well as up to two additional signals, such as invasive blood pressure (ABP), photoplethysmography (PPG), or respiration. Alongside the raw signals, the dataset incorporates metadata generated by the monitor itself. This includes information such as the type of arrhythmia that triggered the alarm (according to the device's automatic criteria), as well as status messages related to the event.

In the original conception of the Challenge, two categories of analysis were delineated:

- In the first event (real time), the algorithm was required to determine whether the alarm was genuine or false, using only the data from the five minutes prior to the

Figure 3. Example of a five-minute recording from the dataset (250 Hz) corresponding to a false asystole alarm



Note: The segment includes four physiological channels: ECG lead II (blue, top panel), ECG lead V (orange), photoplethysmography signal (green), and respiration signal (red). Morphological variability and ICU-related artefacts are evident, illustrating the challenge of discriminating true from false alarms.

trigger.

- In the second event (retrospective), access to 30 seconds following the triggering of the alarm was permitted for the purpose of classification.

This study focuses exclusively on Event 1. This means that the analysed records correspond only to the five minutes prior to the alarm being triggered. There is a unified sampling frequency of 250 Hz, equivalent to 75,000 samples per signal and per record.

In total, the Challenge collected 1,250 segments: 750 were used for training (with public access), and 500 were reserved as a hidden test set. For this study, only the public set of 750 records was accessible, while the hidden set remained unavailable. Each alarm was reviewed by at least two independent experts to establish whether the event corresponded to a true arrhythmia or a false alarm.

The alarms cover five types of clinically relevant event: asystole (ASY), extreme bradycardia (EBR), extreme tachycardia (ETC), ventricular tachycardia (VTA), and ventricular fibrillation/flutter (VFB). Table 1 summarises the distribution of true and false cases in the public training set.

Table 1. Distribution of true and false alarms in the public dataset

Arrhythmia	**True alarm**	**False alarm**	**Total**	**Percentage**
ASY	22	100	122	16.26 %
EBR	46	43	89	11.86 %
ETC	131	9	140	18.66 %
VTA	89	252	341	45.46 %
VFB	6	52	58	7.76 %
Total	**294**	**456**	**750**	**100 %**

Note: Quantities indicate the number of true, false, and total alarms per arrhythmia type. The percentage column represents the proportion of each arrhythmia type within the dataset.

3.2. MODEL CONFIGURATION AND TRAINING SETTINGS

The proposed multimodal Siamese model (Figure 2) was trained using the Adam optimizer for 3000 epochs, with early stopping (patience of 50 epochs) and a *ReduceLROnPlateau* scheduler to ensure stable convergence. The batch size was set to 64, the initial learning rate to 1×10^{-4}, and the L2 regularization coefficient to 0.001. Each input record had a tensor shape of $(c \times n)$, where $c = 4$ represents the number of physiological channels and $n = 75{,}000$ the total number of temporal samples per channel. Two fixed-length segments of shape (4×2500) were extracted from each record, corresponding to the alarm and reference signals, which were then encoded into latent embeddings $\mathbf{e_a}$ and $\mathbf{e_r}$.

The total loss combined binary cross-entropy with a Siamese constraint term, as defined in Eq. (1). For the Zhou Constraint, the weighting factor was set to $w = 1.5$. For the Proposed Constraint, the parameters were defined heuristically as $w = 1 \times 10^{-3}$ and $\beta = 1 \times 10^{2}$, while the margin parameter was fixed to $\alpha = 1$. Dropout with a rate of 0.75 was applied within convolutional blocks, and all weights were initialized using the Xavier uniform method. The learning rate was reduced by a factor of 0.1 when the validation loss stagnated for 15 consecutive epochs.

3.3. MODEL PERFORMANCE VALIDATION

The performance of the proposed model was evaluated using three metrics: the Area Under the Receiver Operating Characteristic Curve (AUC-ROC), the F1-Score, and the official metric defined in the PhysioNet/Computing in Cardiology Challenge 2015. The latter incorporates true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), while introducing a strong penalty on false negatives, as defined in Equation (7).

$$C = \frac{TP + TN}{TP + TN + FP + \mathbf{5} \cdot FN} \tag{7}$$

To ensure robustness and generalization, a stratified five-fold cross-validation was conducted, preserving the distribution of true and false alarms across folds. This approach enabled the evaluation of model behavior under different data partitions. Furthermore, given the limited size of the test set, additional validation was performed by introducing controlled levels of Gaussian noise into the test samples.

The noise addition was implemented by adjusting the power of the perturbation according to a target signal-to-noise ratio (SNR) expressed in decibels (dB). For a clean signal x, the noisy version $\tilde{x}$ is obtained as:

$$\tilde{x} = x + n, \quad n \sim \mathcal{N}\left(0, \sigma^2\right), \tag{8}$$

where the noise variance σ^2 is determined from the signal power P_x and the desired SNR level:

$$\sigma^2 = \frac{P_x}{10^{\frac{\text{SNR}_{\text{dB}}}{10}}}. \tag{9}$$

This procedure allowed evaluation of the model under varying noise intensities, thereby assessing its robustness to signal perturbations commonly encountered in ICU monitoring environments. An illustrative example of how noise addition affects the signals is provided in Appendix 1.

3.4. QUANTITATIVE RESULTS

To assess the contribution of each component within the overall model architecture and to evaluate performance under different configurations, an ablation study was conducted. The study consists of progressively adding elements to the baseline model configuration and observing how each modification impacts overall performance.

Table 2 summarizes the results of this analysis. The first two columns correspond to the signal encoder architecture, indicating whether a simple encoder or the full Siamese network was used as the signal-processing module. The next two columns reflect the inclusion of additional inputs beyond the raw physiological signals, namely the one-hot encoded arrhythmia type and the output of the rule-based model. The following three columns report the type of loss functions and constraints applied (binary cross-entropy, Zhou Loss, or the proposed Custom Loss). Data augmentation was used throughout all configurations. Two performance metrics are reported: the official Challenge score and the F1-Score, each expressed as the mean $\pm$ standard deviation across five-fold cross-validation. Each row represents a distinct experimental configuration, and a complete version of the ablation study is provided in Appendix 2.

Table 2. Ablation study of the proposed model showing the effect of each module and training configuration.

Model Configuration				**Loss**			**Metrics (std)**	
Encoder	**Siamese**	**Arrhythmia**	**Rule-Based**	**BCE**	**Zhou Loss**	**Custom Loss**	**Chall. Score**	**F1-Score**
✓				✓			0.6083 (0.0363)	0.7591 (0.0305)
	✓			✓			0.5831 (0.0820)	0.7526 (0.0441)
	✓				✓		0.6137 (0.0354)	0.7517 (0.0487)
	✓					✓	0.6453 (0.0347)	0.7534 (0.0450)
	✓	✓		✓			0.6608 (0.0791)	0.8120 (0.0759)
	✓	✓			✓		0.7407 (0.0745)	0.8280 (0.0749)
	✓	✓				✓	0.7351 (0.0481)	0.8286 (0.0396)
	✓	✓	✓	✓			0.8216 (0.0747)	0.8988 (0.0883)
	✓	✓	✓		✓		0.8267 (0.0619)	0.9042 (0.0501)
	✓	✓	✓			✓	**0.8498 (0.0686)**	**0.9101 (0.0406)**
			✓				0.8088 (0.0546)	0.8690 (0.0576)
Replied Zhou Model					✓		0.8348 (0.0652)	0.9005 (0.0612)

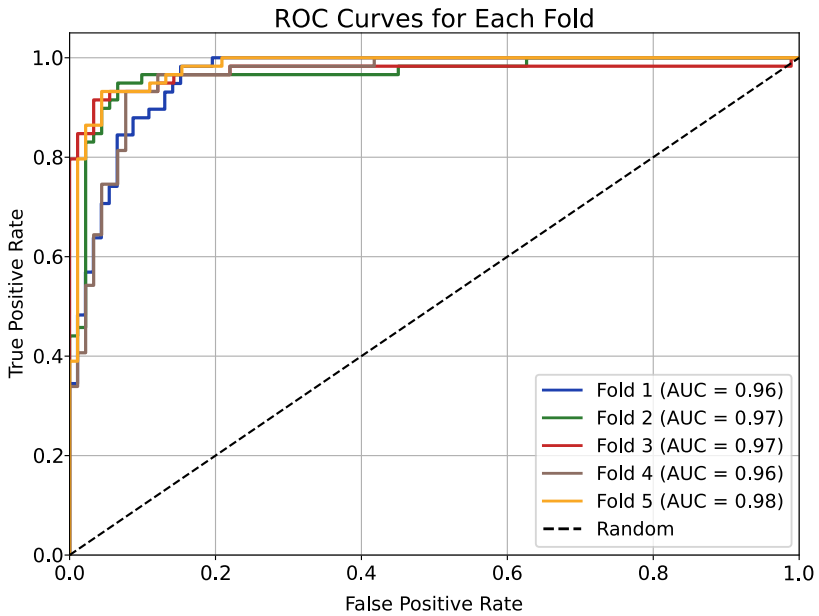
Ablation study of the proposed model across different configurations. Each horizontal block corresponds to a distinct model configuration, where components such as the Siamese encoder, arrhythmia metadata, and rule-based features were selectively enabled. All experiments were conducted with data augmentation. The results, reported as mean $\pm$ standard deviation over five-fold cross-validation, demonstrate the effect of different loss functions (BCE, Zhou Loss, and Custom Loss). The bold entry highlights the best performance achieved by the proposed model, and the final row presents the results obtained by replicating the model proposed by Zhou et al. [45].

Horizontal separators in the table indicate changes in the underlying model architecture. Within each architecture, multiple experiments were performed by varying

the loss functions and constraints. The table also includes the results obtained by replicating the model proposed in Zhou et al., using only the information provided in their publication.

All experiments in the ablation study were conducted using the same random seed (42), which was used to generate the train, validation, and test splits, ensuring full comparability of results across configurations. The best performance was achieved with the proposed model, reaching an official Challenge score of 0.8498 ± 0.0686 and an F1-score of 0.9101 ± 0.0406. These results demonstrate the effectiveness of the proposed architecture for detecting false alarms associated with cardiac arrhythmias. The training and validation behavior of the best-performing model across the five-fold cross-validation is shown in Appendix 3, while the corresponding ROC curves for each fold are presented in Figure 4. The model achieved consistent convergence in all folds, with AUC values ranging from 0.96 to 0.98.

Figure 4. Receiver Operating Characteristic (ROC) curves for each fold of the five-fold cross-validation using the best-performing model



The Area Under the Curve (AUC) is reported for each fold.

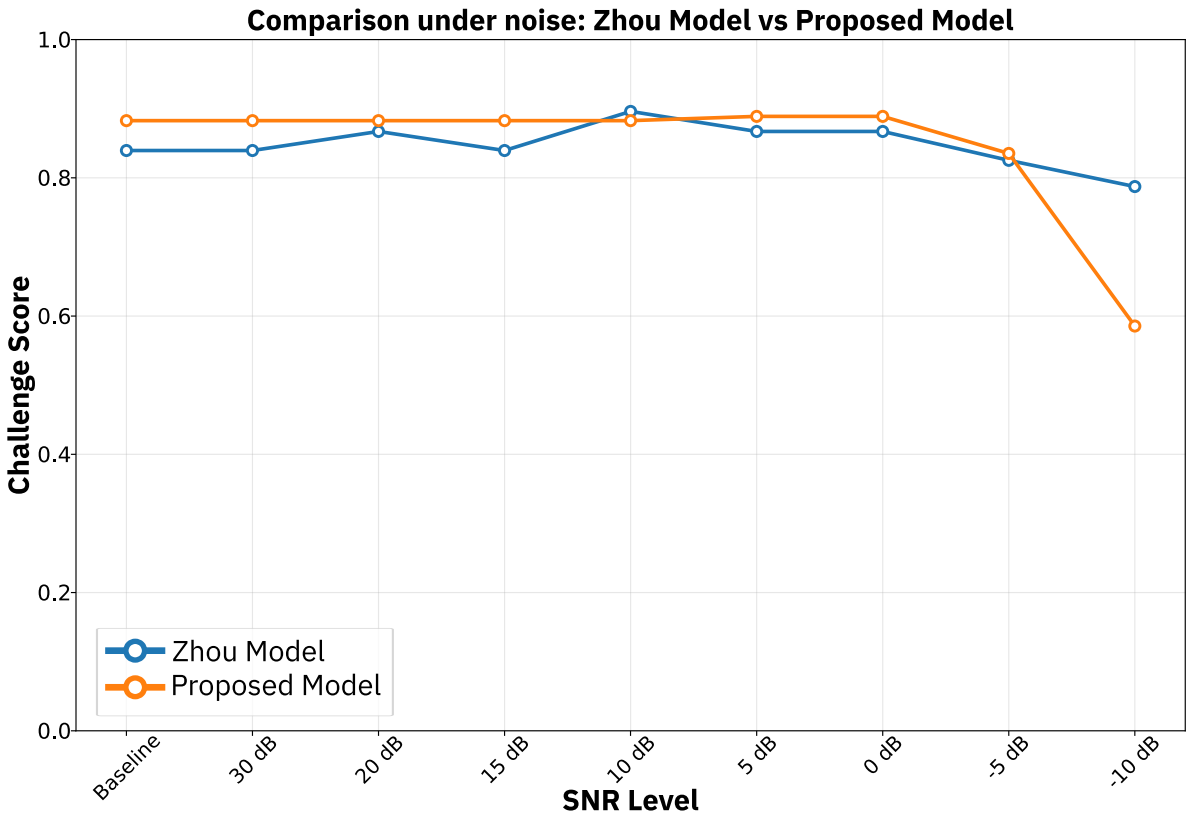
To evaluate the generalizability of the proposed model, a controlled experiment was conducted in which 600 samples were allocated for training and 150 for testing. The model was trained on the training set and subsequently evaluated on the test set under progressively increasing noise levels introduced through the technique described in Section 3.3. Specifically, the signals were corrupted with Gaussian noise at signal-to-noise ratio (SNR) levels of 30, 20, 15, 10, 5, 0, -5, and -10 dB. The same evaluation procedure was also applied to the replicated model from Zhou et al.

Figure 5 presents the evolution of the official Challenge score as the SNR decreases. Both models maintained stable performance down to 0 dB, with the proposed model consistently outperforming the replicated Zhou model. However, at more severe noise levels (-5 and -10 dB), the replicated model exhibited better robustness and preserved higher performance compared to the proposed model.

Appendix 4 reports the evolution of the official Challenge score and the F1-score across different SNR levels added to the test signals. Both metrics follow a similar trend as noise increases, although the F1-score exhibits a more stable behavior and does not degrade as sharply as the Challenge score under high noise conditions.

The final architecture contained a total of 878,925 trainable parameters. All experiments were conducted on a computing system equipped with an NVIDIA Tesla V100 GPU (16 GB) and a dual Intel Xeon Gold 6130 CPU. The average inference time per sample was 0.020 seconds, demonstrating the feasibility of deploying the model in real-time ICU monitoring scenarios.

Figure 5. Comparison of the proposed model and the replicated Zhou model under different SNR conditions



4. DISCUSSION

The primary objective of this study was to evaluate the feasibility of integrating rule-based models with machine learning algorithms in order to improve the detection of false alarms in cardiac arrhythmia monitoring systems in intensive care units (ICUs) settings. Overall, the experimental results support this objective, demonstrating that the proposed model effectively integrates heterogeneous physiological signals with arrhythmia metadata and rule-based features to generate discriminative multimodal representations. The findings consistently show that the proposed framework improves upon standard baselines and surpasses the replicated state-of-the-art model used as reference, confirming the relevance and validity of the methodological contributions.

Interpretation of Findings

A first set of insights emerges from the ablation study presented in Table 2, which evaluates the contribution of each module of the architecture. The experiments reveal that data augmentation played a consistent and beneficial role across multiple configurations. This improvement is largely attributable to the temporal variability induced during pair construction: while the alarm window remains fixed, the reference window is sampled from a wide temporal interval, enabling the generation of highly diverse signal pairs. This variability helps the model generalize better to unseen examples, particularly in scenarios where noise, artifacts, and physiological variability are commonplace.

A second critical observation concerns the integration of arrhythmia-type information. The one-hot arrhythmia vector, when processed through an embedding layer, systematically enhanced performance. This behavior suggests that the embedding

layer successfully learns associations between the arrhythmia category and the latent physiological patterns extracted by the encoders, allowing the model to adjust its internal representation depending on the alarm type.

Moreover, the inclusion of rule-based outputs proved to be essential for reaching the best overall results. Because the rule-based system encodes domain knowledge specific to each arrhythmia, its concatenation with the arrhythmia embedding introduces a structured and interpretable prior into the learning process. The model therefore benefits from complementary information: the rule-based scores provide expert-defined cues, while the embeddings allow the network to capture statistical regularities from the raw data. Together, these components significantly enhance the system's ability to distinguish true alarms from false ones.

Finally, the ablation study highlights that the Custom Loss proposed in this work generally outperforms the Zhou Loss. This improvement can be explained by examining their mathematical formulations. The Zhou constraint relies on the inner product $\mathbf{e_a^T} \cdot \mathbf{e_r}$ between non-negative feature vectors obtained after a ReLU activation. As a consequence, the similarity scores are restricted to non-negative values, reducing discriminative power, particularly for samples in which negative correlations between embeddings would be informative. In addition, the formulation of the constraint imposes asymmetric penalties for the positive and negative classes, which may limit numerical stability.

By contrast, the Custom Loss introduced in Equation (5) leverages the normalized Euclidean distance between embeddings, enabling the model to meaningfully represent both similarity and dissimilarity in a continuous and unconstrained metric space. The inclusion of a margin α enforces a minimum separation for positive (true alarm) pairs, while the weighting factor β, inspired by focal loss, strengthens the penalization of hard-to-classify positive examples. This formulation improves stability, enhances discriminative capacity, and ultimately leads to superior performance across the

ablation configurations.

Comparison with Previous Work

Table 3 presents a direct comparison between the proposed model and the method by Zhou et al., one of the leading approaches in the field. The table includes both the original results reported by the authors and the performance obtained when replicating their model based on the information available in their publication.

Table 3. Comparison of the proposed model with the state-of-the-art model by Zhou et al.

Models	Score	F1 - Score
Replied Zhou Model	0.8348 (0.0652)	0.9005 (0.0612)
Zhou Model	0.8700 (0.0484)	————
Proposed Model	0.8498 (0.0686)	0.9101 (0.0406)

[46].

The proposed model achieves a Challenge score of 0.8498 ± 0.0686, surpassing the replicated Zhou model, which obtains 0.8348 ± 0.0652. As described in Section 3, the proposed architecture contains 878,925 trainable parameters, whereas the replicated Zhou model uses 3,282,753 parameters. This difference suggests that the proposed design extracts more informative representations while requiring substantially fewer parameters. The resulting model is therefore more lightweight and better suited for real-time or resource-constrained deployment scenarios. This interpretation is further supported by the average inference time of 0.020 seconds per sample, which confirms that the model is computationally efficient and can be integrated into continuous monitoring systems.

However, the proposed model does not reach the performance reported by Zhou et al. in their original publication, where they achieved a Challenge score of 0.8700 ± 0.0484. Because the authors did not provide their implementation or codebase, it is plausible that subtle architectural choices, preprocessing decisions, or training heuristics used in their original setup may not have been fully reproduced. This discre-

pancy highlights the sensitivity of deep-learning-based pipelines to implementation details and reinforces the importance of reproducible research practices in the field. Nevertheless, the improvements observed when comparing the proposed model with the replicated baseline indicate that the methodological contributions presented here—particularly the use of multimodal fusion, arrhythmia embeddings, rule-based priors, and the Custom Loss—represent meaningful advances in designing architectures for false alarm reduction.

Implications

The findings of this study have several implications. From a methodological perspective, they demonstrate that combining heterogeneous data sources—raw physiological signals, categorical arrhythmia metadata, and structured rule-based outputs—can significantly enhance discriminative performance in Siamese neural architectures. This strategy may be applicable beyond false alarm reduction, potentially benefiting tasks such as arrhythmia classification with reduced leads, multimodal patient monitoring, and real-time clinical decision support.

From a practical standpoint, the lightweight nature of the proposed architecture suggests that it could be deployed in resource-limited ICU environments, embedded systems, or bedside monitoring devices. The model's robustness under multiple noise conditions further supports its relevance for real-world settings, where measurements often suffer from artifacts, motion-related disturbances, and variable signal quality.

Limitations

Although the results are promising, several aspects merit careful consideration. The experiments were conducted using the Challenge database, which, while representative, does not fully capture the full spectrum of physiological variability encountered

across different hospitals, patient populations, or acquisition devices. As a result, the generalization of the model to broader clinical settings remains to be validated.
Additionally, while the replication of the Zhou model was performed as faithfully as possible, the absence of the original implementation introduces inherent uncertainty in the comparison. Certain preprocessing details, architectural nuances, or hyperparameter choices may differ from the original authors' setup, potentially influencing relative performance. Nevertheless, these discrepancies emphasize the value of transparent reporting and reproducible methodologies in algorithmic research.
Finally, although the proposed Custom Loss demonstrated strong empirical behavior, its formulation includes hyperparameters that may need tuning for other datasets or tasks. Further analyses could be performed to explore its behavior under different regimes of class imbalance or in architectures beyond the Siamese framework.

Future Work

Future research could build on these findings in several directions. First, validating the model across multi-center datasets or through prospective studies would provide stronger evidence of clinical utility. Second, the rule-based component could be expanded to incorporate additional expert knowledge or dynamically adapt its thresholds based on patient-specific parameters. Third, extending the multimodal fusion framework to include additional data sources—such as demographic variables, clinical notes, or complementary biosignals—could further enhance performance. Finally, the Custom Loss could be explored in applications such as contrastive learning, representation learning, or multimodal retrieval tasks, potentially revealing broader utility beyond the scope of false alarm detection.

5. CONCLUSIONS

The multimodal Siamese framework developed in this study provides an effective and computationally efficient solution for reducing false arrhythmia alarms in ICUs. By integrating raw physiological signals, arrhythmia-type embeddings, and rule-based outputs, the model captures complementary information that significantly enhances its ability to discriminate between true and false alarms. The combination of convolutional encoders, a targeted data augmentation strategy, and a proposed loss function contributed decisively to the robustness and generalization of the system.

The experimental results confirm the impact of this design: the proposed approach achieved a Challenge score of **0.8498** $\pm$ **0.0686** and an F1-score of **0.9101** $\pm$ **0.0406**, outperforming a faithfully replicated state-of-the-art baseline while requiring nearly three times fewer parameters. These findings indicate that lightweight architectures, when combined with structured priors and modality-aware optimization, can offer substantial benefits for real-time ICU monitoring.

Despite these strengths, the evaluation was limited to a single public dataset, which may not fully reflect the variability encountered across hospitals and acquisition devices. Broader validation is therefore necessary to confirm clinical applicability. Future work should explore multi-center datasets, extend the rule-based component with additional clinical knowledge, and investigate the integration of other data sources such as patient metadata or complementary biosignals. Further refinement of the proposed loss function may also reveal its utility in related multimodal or contrastive learning tasks.

Overall, the results demonstrate that hybrid approaches combining deep learning with expert knowledge can enhance interpretability, stability, and performance in critical-care monitoring systems, offering a promising direction for future research and deployment.

BIBLIOGRAPHY

Aboukhalil, Anton et al. «Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform». En: *Journal of Biomedical Informatics* 41 (3 jun. de 2008), págs. 442-451. DOI: 10.1016/J.JBI.2008.03.003 (vid. pág. 14).

Afghah, Fatemeh, Abolfazl Razi y Kayvan Najarian. *A Shapley Value Solution to Game Theoretic-based Feature Reduction in False Alarm Detection*. 2015. arXiv: 1512.01680 [cs.CV].

Ansari, Sardar, Ashwin Belle y Kayvan Najarian. «Multi-modal integrated approach towards reducing false arrhythmia alarms during continuous patient monitoring: The Physionet Challenge 2015». En: *Computing in Cardiology* 42 (feb. de 2015). DOI: 10.1109/CIC.2015.7411127 (vid. pág. 12).

Ansari, Sardar et al. «Suppression of false arrhythmia alarms in the ICU: a machine learning approach». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1186. DOI: 10.1088/0967-3334/37/8/1186 (vid. pág. 16).

Antink, Christoph Hoog y Steffen Leonhardt. «Reducing false arrhythmia alarms using robust interval estimation and machine learning». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 285-288. DOI: 10.1109/CIC.2015.7408642.

Antink, Christoph Hoog, Steffen Leonhardt y Marian Walter. «Reducing false alarms in the ICU by quantifying self-similarity of multimodal biosignals». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1233. DOI: 10.1088/0967-3334/37/8/1233.

Ardila, Karol et al. «Detecting False Arrhythmias Alarms in the ICU Using a Deep Learning Approach». En: *2024 XXIV Symposium of Image, Signal Processing, and Artificial Vision (STSIVA)*. 2024, págs. 1-5. DOI: 10.1109/STSIVA63281.2024.10637887 (vid. pág. 17).

Au-Yeung, Wan Tai M. et al. «Real-time machine learning-based intensive care unit alarm classification without prior knowledge of the underlying rhythm». En: *European Heart Journal. Digital Health* 2 (3 sep. de 2021), pág. 437. DOI: 10.1093/EHJDH/ZTAB058.

Au-Yeung, Wan Tai M. et al. «Reduction of false alarms in the intensive care unit using an optimized machine learning based approach». En: *npj Digital Medicine* 2 (1 dic. de 2019), págs. 1-5. DOI: 10.1038/S41746-019-0160-7;SUBJMETA (vid. págs. 17, 18).

Baumgartner, Benedikt, Kolja Rodel y Alois Knoll. «A data mining approach to reduce the false alarm rate of patient monitors». En: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2012), págs. 5935-5938. DOI: 10.1109/EMBC.2012.6347345 (vid. pág. 12).

Behar, Joachim et al. «ECG signal quality during arrhythmia and its application to false alarm reduction». En: *IEEE Transactions on Biomedical Engineering* 60 (6 2013), págs. 1660-1666. DOI: 10.1109/TBME.2013.2240452 (vid. pág. 14).

Bitan, Yuval y Michael F. O'Connor. «Correlating data from different sensors to increase the positive predictive value of alarms: an empiric assessment». En: *F1000 Research 2012 1:45* 1 (nov. de 2012), pág. 45. DOI: 10.12688/f1000research.1-45.v1 (vid. pág. 14).

Bollepalli, Sandeep Chandra et al. «Real-Time Arrhythmia Detection Using Hybrid Convolutional Neural Networks». En: *Journal of the American Heart Association* 10 (23 dic. de 2021), pág. 23222. DOI: 10.1161/JAHA.121.023222 (vid. pág. 18).

Boynton, Jack y Byung Suk Lee. «Deep Learning Based Classification of True/False Arrhythmia Alarms in the Intensive Care Unit». En: *Computing in Cardiology* 2021-September (2021). DOI: 10.23919/CINC53138.2021.9662874 (vid. pág. 17).

Caballero, Miguel y Grace M. Mirsky. «Reduction of false cardiac arrhythmia alarms through the use of machine learning techniques». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 1169-1172. DOI: 10.1109/CIC.2015.7411124 (vid. pág. 12).

Charbonnier, S., D. Pean y S. Gentil. «Alarm Filtering in Intensive Care Units Using Multivariable Analysis of Physiological Parameters». En: *Fault Detection, Supervision and Safety of Technical Processes 2006* 2 (ene. de 2007), págs. 1127-1132. DOI: 10.1016/B978-008044485-7/50190-1.

Clifford, Gari D. et al. «The PhysioNet/Computing in Cardiology Challenge 2015: Reducing false arrhythmia alarms in the ICU». En: *Computing in Cardiology* 42 (2015). Cited by: 116; All Open Access, Green Open Access, 273 – 276. DOI: 10.1109/CIC.2015.7408639 (vid. págs. 15, 30).

Clifford, Gari D et al. «Using the blood pressure waveform to reduce critical false ECG alarms | IEEE Conference Publication | IEEE Xplore». En: *2006 Computers in Cardiology*. IEEE, 2006, págs. 829-832 (vid. pág. 14).

Couto, Paula, Ruben Ramalho y Rui Rodrigues. «Suppression of false arrhythmia alarms using ECG and pulsatile waveforms». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 749-752. DOI: 10.1109/CIC.2015.7411019.

Daluwatte, Chathuri et al. «Heartbeat fusion algorithm to reduce false alarms for arrhythmias». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 745-748. DOI: 10.1109/CIC.2015.7411018.

Drew, Barbara J. et al. «Insights into the Problem of Alarm Fatigue with Physiologic Monitor Devices: A Comprehensive Observational Study of Consecutive Intensive Care Unit Patients». En: *PLOS ONE* 9 (10 oct. de 2014), e110274. DOI: 10.1371/JOURNAL.PONE.0110274 (vid. pág. 10).

Drew, Barbara J. et al. «Practice Standards for Electrocardiographic Monitoring in Hospital Settings». En: *Circulation* 110 (17 oct. de 2004), págs. 2721-2746. DOI: 10.1161/01.CIR.0000145144.56673.59 (vid. pág. 10).

Eerikainen, Linda M. et al. «Reduction of false arrhythmia alarms using signal selection and machine learning». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1204. DOI: 10.1088/0967-3334/37/8/1204.

Eerikäinen, Linda M. et al. «Decreasing the false alarm rate of arrhythmias in intensive care using a machine learning approach». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 293-296. DOI: 10.1109/CIC.2015.7408644.

— «Decreasing the false alarm rate of arrhythmias in intensive care using a machine learning approach». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 293-296. DOI: 10.1109/CIC.2015.7408644.

Elghazzawi, Ziad y Frederick Geheb. «A knowledge-based system for arrhythmia detection». En: *Computers in Cardiology* 0 (0 1996), págs. 541-544. DOI: 10.1109/CIC.1996.542593.

Fallet, Sibylle, Sasan Yazdani y Jean Marc Vesin. «A multimodal approach to reduce false arrhythmia alarms in the intensive care unit». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 277-280. DOI: 10.1109/CIC.2015.7408640 (vid. págs. 11, 15).

— «False arrhythmia alarms reduction in the intensive care unit: a multimodal approach». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1217. DOI: 10.1088/0967-3334/37/8/1217 (vid. págs. 11, 16, 18).

Fallet, Sibylle et al. «Reducing false alarms in the ICU by quantifying self-similarity of multimodal biosignals». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1233. DOI: 10.1088/0967-3334/37/8/1233.

Gajowniczek, Krzysztof, Iga Grzegorczyk y Tomasz Ząbkowski. «Reducing False Arrhythmia Alarms Using Different Methods of Probability and Class Assignment in Random Forest Learning Methods». En: *Sensors 2019, Vol. 19, Page 1588* 19 (7 abr. de 2019), pág. 1588. DOI: 10.3390/S19071588 (vid. pág. 17).

Gamboa, John Cristian Borges. *Deep Learning for Time-Series Analysis*. 2017. arXiv: 1701.01887 [cs.LG].

He, Runnan et al. «Reducing false arrhythmia alarms in the ICU using novel signal quality indices assessment method». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 1189-1192. DOI: 10.1109/CIC.2015.7411129 (vid. págs. 11, 16).

Hever, Gal et al. «Machine learning applied to multi-sensor information to reduce false alarm rate in the ICU». En: *Journal of Clinical Monitoring and Computing* 34 (2 abr. de 2020), págs. 339-352. DOI: 10.1007/S10877-019-00307-X/TABLES/5.

Hong, Shenda et al. «ENCASE: An ENsemble ClASsifiEr for ECG classification using expert features and deep neural networks». En: *Computing in Cardiology* 44 (2017), págs. 1-4. DOI: 10.22489/CINC.2017.178-245 (vid. pág. 13).

Hyvärinen, Aapo e Hiroshi Morioka. «Unsupervised Feature Extraction by Time-Contrastive Learning and Nonlinear ICA». En: *Advances in Neural Information Processing Systems* (mayo de 2016), págs. 3772-3780.

Kalidas, V. y L. S. Tamil. «Cardiac arrhythmia classification using multi-modal signal analysis». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1253. DOI: 10.1088/0967-3334/37/8/1253 (vid. pág. 12).

Kiyasseh, Dani, Tingting Zhu y David A. Clifton. «CLOCS: Contrastive Learning of Cardiac Signals Across Space, Time, and Patients». En: *Proceedings of Machine Learning Research* 139 (mayo de 2020), págs. 5606-5615.

Krasteva, Vessela et al. «Real-time arrhythmia detection with supplementary ECG quality and pulse wave monitoring for the reduction of false alarms in ICUs». En: *Physiological measurement* 37 (8 jul. de 2016), págs. 1273-1297. DOI: 10.1088/0967-3334/37/8/1273 (vid. pág. 16).

Lameski, Petre et al. «Suppression of Intensive Care Unit False Alarms Based on the Arterial Blood Pressure Signal». En: *IEEE Access* 5 (2017), págs. 5829-5836. DOI: 10.1109/ACCESS.2017.2690380 (vid. pág. 17).

Lehman, Eric P. et al. «Representation Learning Approaches to Detect False Arrhythmia Alarms from ECG Dynamics». En: *Proceedings of machine learning research* 85 (2018), pág. 571 (vid. pág. 17).

Li, Qiao y Gari D. Clifford. «Signal quality and data fusion for false alarm reduction in the intensive care unit». En: *Journal of Electrocardiology* 45 (6 nov. de 2012), págs. 596-603. DOI: 10.1016/J.JELECTROCARD.2012.07.015 (vid. pág. 14).

Liu, Chengyu, Lina Zhao y Hong Tang. «Reduction of False Alarms in Intensive Care Unit using Multi-feature Fusion Method». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 741-744. DOI: 10.1109/CIC.2015.7411017 (vid. pág. 15).

Manna, Tishya, Aleena Swetapadma y Moloud Abdar. «Decision Tree Predictive Learner-Based Approach for False Alarm Detection in ICU». En: *Journal of Medical Systems* 43 (7 jul. de 2019), págs. 1-13. DOI: 10.1007/S10916-019-1337-Y/TABLES/5 (vid. pág. 12).

Marshall, John C. et al. «What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine». En: *Journal of critical care* 37 (feb. de 2017), págs. 270-276. DOI: 10.1016/J.JCRC.2016.07.015 (vid. pág. 10).

Morrison, Wynne E. et al. *Noise, stress, and annoyance in a pediatric intensive care units : Critical Care Medicine*. 2003 (vid. pág. 11).

Mousavi, Sajad, Atiyeh Fotoohinasab y Fatemeh Afghah. «Single-modal and multi-modal false arrhythmia alarm reduction using attention-based convolutional and recurrent neural networks». En: *PLOS ONE* 15 (1 ene. de 2020), e0226990. DOI: 10.1371/JOURNAL.PONE.0226990 (vid. pág. 17).

Onega, Tracy et al. «Breast cancer screening in an era of personalized regimens: A conceptual model and National Cancer Institute initiative for risk-based and preference-based approaches at a population level». En: *Cancer* 120.19 (2014), págs. 2955-2964.

Oroojeni, M. J. Hooman, Mohammad Majid Al-Rifaie y Mihalis A. Nicolaou. «Deep neuroevolution: Training deep neural networks for false alarm detection in intensive care units». En: *European Signal Processing Conference* 2018-September (nov. de 2018), págs. 1157-1161. DOI: 10.23919/EUSIPCO.2018.8552944.

Parthasarathy, Sairam y Martin J. Tobin. «Sleep in the intensive care unit». En: *Intensive Care Medicine* 30 (2 feb. de 2004), págs. 197-206. DOI: 10.1007/S00134-003-2030-6/FIGURES/4 (vid. pág. 11).

Pei, Wenjie, David M. J. Tax y Laurens van der Maaten. *Modeling Time Series Similarity with Siamese Recurrent Networks*. 2016. arXiv: 1603.04713 [cs.CV].

Plesinger, F. et al. «Taming of the monitors: reducing false alarms in intensive care units». En: *Physiological Measurement* 37 (8 jul. de 2016), pág. 1313. DOI: 10.1088/0967-3334/37/8/1313 (vid. pág. 16).

Plesinger, Filip et al. «False alarms in intensive care unit monitors: Detection of life-threatening arrhythmias using elementary algebra, descriptive statistics and fuzzy logic». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 281-284. DOI: 10.1109/CIC.2015.7408641 (vid. págs. 15, 16, 23).

Rodriguez, Harold H., Hans Y. Garcia y Carlos A. Fajardo. «Embedding-Enhanced Multimodal Siamese Network for False Arrhythmia Alarm Reduction in Intensive Care Units». En: *2025 25th Symposium of Image, Signal Processing, and Artificial Vision, STSIVA 2025* (2025). DOI: 10.1109/STSIVA66383.2025.11156764 (vid. pág. 18).

Sadr, Nadi et al. «Reducing false arrhythmia alarms in the ICU by Hilbert QRS detection». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 1173-1176. DOI: 10.1109/CIC.2015.7411125 (vid. pág. 16).

Teo, Soo Kng et al. «Reducing false arrhythmia alarms in the ICU». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 1177-1180. DOI: 10.1109/CIC.2015.7411126.

Tsien, Christine L y James C Fackler. *Poor prognosis for existing monitors in the intensive care units : Critical Care Medicine*. 1997 (vid. pág. 11).

Tsimenidis, Charalampos y Alan Murray. «Reliability of clinical alarm detection in intensive care units». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 1185-1188. DOI: 10.1109/CIC.2015.7411128 (vid. pág. 11).

Wang, J., C. L. Yeo y A. Aguirre. «Design and evaluation of a new multi-lead arrhythmia monitoring algorithm». En: *Computers in Cardiology* (1999), págs. 675-678. DOI: 10.1109/CIC.1999.826061.

Wu, Feng et al. *A Diffusion Model with Contrastive Learning for ICU False Arrhythmia Alarm Reduction*. 2023 (vid. pág. 18).

Yanar, Erdem y Yesim Serinagaoglu Dogrusoz. «False Ventricular-Fibrillation/Flutter Alarm Reduction of Patient Monitoring Systems in Intensive Care Units». En: *MeMeA 2018 - 2018 IEEE International Symposium on Medical Measurements and Applications, Proceedings* (ago. de 2018). DOI: 10.1109/MEMEA.2018.8438601 (vid. pág. 11).

Yu, Qiang et al. «Intensive Care Unit False Alarm Identification Based on Convolution Neural Network». En: *IEEE Access* 9 (2021), págs. 81841-81854. DOI: 10.1109/ACCESS.2021.3086862 (vid. págs. 12, 17).

Zaeri-Amirani, Mohammad, Fatemeh Afghah y Sajad Mousavi. «A Feature Selection Method Based on Shapley Value to False Alarm Reduction in ICUs, A Genetic-Algorithm Approach». En: *Proceedings of the Annual International Conference of the*

IEEE Engineering in Medicine and Biology Society, EMBS 2018-July (abr. de 2018), págs. 319-323. DOI: 10.1109/EMBC.2018.8512266.

Zhang, Qiang et al. «Reducing false arrhythmia alarm rates using robust heart rate estimation and cost-sensitive support vector machines». En: *Physiological Measurement* 38 (2 ene. de 2017), pág. 259. DOI: 10.1088/1361-6579/38/2/259 (vid. pág. 16).

Zhou, Yuerong et al. «A contrastive learning approach for ICU false arrhythmia alarm reduction». En: *Scientific Reports 2022 12:1* 12 (1 mar. de 2022), págs. 1-10. DOI: 10.1038/s41598-022-07761-9 (vid. págs. 18, 19, 24, 28, 35, 41, 56).

Zihlmann, Martin, Dmytro Perekrestenko y Michael Tschannen. «Convolutional recurrent neural networks for electrocardiogram classification». En: *Computing in Cardiology* 44 (2017), págs. 1-4. DOI: 10.22489/CINC.2017.070-060 (vid. pág. 12).

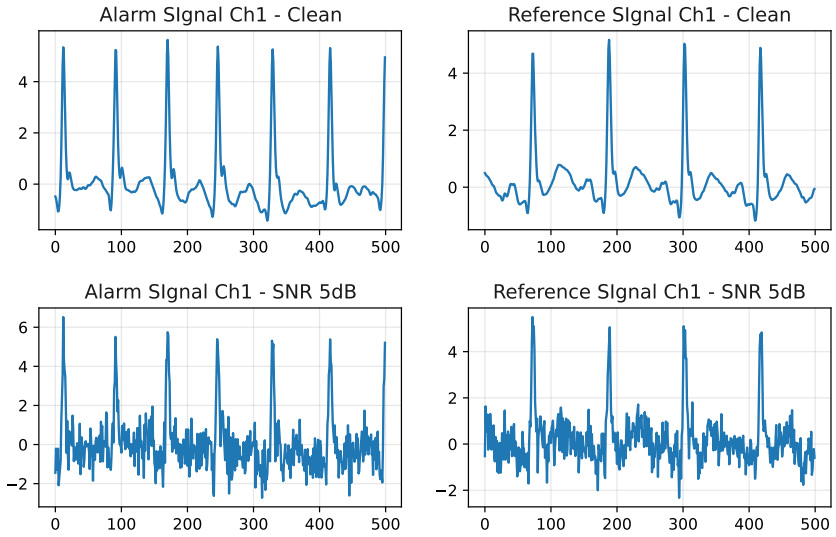
Zong, Wei. «Reduction of false critical ECG alarms using waveform features of arterial blood pressure and/or photoplethysmogram signals». En: *Computing in Cardiology* 42 (feb. de 2015), págs. 289-292. DOI: 10.1109/CIC.2015.7408643 (vid. pág. 15).

ANNEXES

Annex A. Noise Addition

This appendix illustrates the process used to add Gaussian noise to the signals. Figure 6 shows how noise addition affects both the alarm and reference signals. During the study, multiple noise levels were applied to evaluate the model's robustness under varying signal quality conditions.

Figure 6. Example of noise addition in one physiological channel for both the alarm and reference signals. The top panels show the clean signals, while the bottom panels illustrate the corresponding signals after the addition of Gaussian noise at 5 dB SNR. This visualization highlights how controlled perturbations affect waveform morphology and amplitude, emulating realistic noise conditions in ICU environments.



Annex B. Full ablation study

This appendix provides a comprehensive overview of the ablation study conducted in this research. The study presents the complete set of experimental results obtained both with and without data augmentation. As shown in Table 4, the effect of data augmentation becomes more evident, demonstrating that in most cases it provides valuable information to the model, enabling improved learning and overall performan-

ce.

Table 4. Full Ablation study of the proposed model.

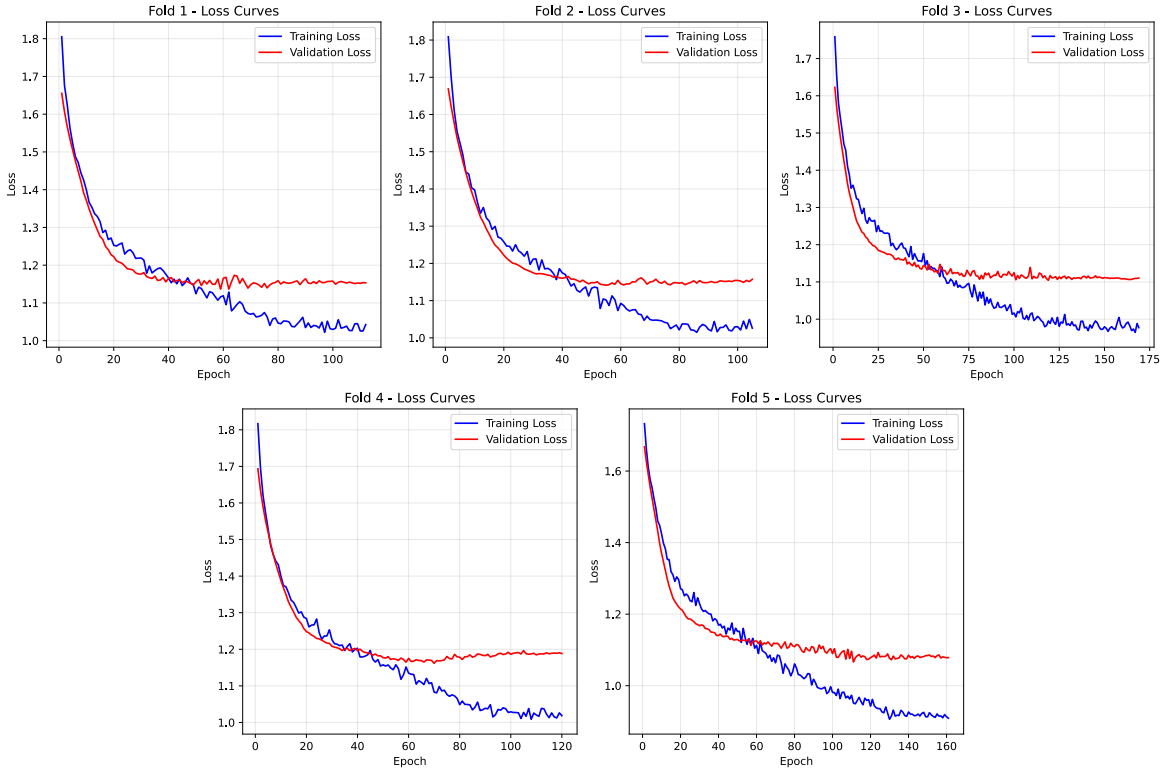
Model Configuration				Loss			Augmentation	Metrics (± std)	
Encoder	Siamese	Arrhythmia	Rule-Based	BCE	Zhou Loss	Custom Loss		Score	F1-score
✓				✓			No	0.6035 (0.0243)	0.7493 (0.0166)
✓				✓			Yes	0.6083 (0.0363)	0.7591 (0.0305)
	✓			✓			No	0.5888 (0.0857)	0.7488 (0.0584)
	✓			✓			Yes	0.5831 (0.0820)	0.7526 (0.0441)
	✓				✓		No	0.5901 (0.0547)	0.7513 (0.0513)
	✓				✓		Yes	0.6137 (0.0354)	0.7517 (0.0487)
	✓					✓	No	0.6106 (0.0391)	0.7494 (0.0318)
	✓					✓	Yes	0.6453 (0.0347)	0.7534 (0.0450)
	✓	✓		✓			No	0.6462 (0.0403)	0.8072 (0.0432)
	✓	✓		✓			Yes	0.6608 (0.0791)	0.8120 (0.0759)
	✓	✓			✓		No	0.7110 (0.0634)	0.8167 (0.0490)
	✓	✓			✓		Yes	0.7407 (0.0745)	0.8280 (0.0749)
	✓	✓				✓	No	0.7153 (0.0741)	0.8205 (0.0611)
	✓	✓				✓	Yes	0.7351 (0.0481)	0.8286 (0.0396)
	✓	✓	✓	✓			No	0.8163 (0.0576)	0.8994 (0.0567)
	✓	✓	✓	✓			Yes	0.8216 (0.0747)	0.8988 (0.0883)
	✓	✓	✓		✓		No	0.8162 (0.0488)	0.8930 (0.0543)
	✓	✓	✓		✓		Yes	0.8267 (0.0619)	0.9042 (0.0501)
	✓	✓	✓			✓	No	0.8309 (0.0546)	0.9028 (0.0513)
	✓	✓	✓			✓	Yes	**0.8498 (0.0686)**	**0.9101 (0.0406)**
			✓				No	0.8088 (0.0546)	0.8690 (0.0576)
Replied Zhou Model					✓		No	0.8301 (0.0652)	0.8999 (0.0412)
Replied Zhou Model					✓		Yes	0.8348 (0.0652)	0.9005 (0.0612)
Zhou Model							Yes	**0.8700 (0.0484)**	-.— (-.—)

Ablation study of the proposed model across different configurations. Each horizontal block corresponds to a distinct model configuration, where components such as the Siamese encoder, arrhythmia metadata, and rule-based features were selectively enabled. The experiments also evaluate the effect of different loss functions (BCE, Zhou Loss, and Custom Loss) and the use of data augmentation. Results are reported as mean ± standard deviation over five-fold cross-validation. The first bold entry highlights the best performance achieved in this study, while the second bold entry corresponds to the performance reported by Zhou et al [47].

Annex C. Training and Validation Loss Curves

This appendix presents Figure 7, which shows the training and validation loss curves for each of the five folds during cross-validation. These curves correspond to the proposed model under its complete configuration, providing a detailed view of the convergence behavior and training stability across all folds. Across the five folds, a consistent training pattern can be observed, where both training and validation losses decrease rapidly during the initial epochs. However, approximately between epochs 40 and 50, the validation loss tends to stagnate or slightly increase, while the training loss continues to decrease. This divergence indicates the onset of overfitting, which is consistent across folds. A plausible explanation for this behavior is that, at this stage, the optimization process begins to extract increasingly fine-grained patterns from the augmented training data that are no longer generalizable to the validation set. As a result, while the model continues to improve its fit to the training data—benefiting from the applied data augmentation—the learned representations provide diminishing gains in generalization performance.

Figure 7. Training and validation loss curves for the proposed model across the five folds of cross-validation. The figure illustrates consistent convergence behavior, with the onset of overfitting observed around epochs 40–50, where the validation loss stagnates while the training loss continues to decrease.



Annex D. Performance under different noise levels (SNR)

Table 5 reports the evolution of the official Challenge score and the F1-score across different SNR levels added to the test signals. Both metrics follow a similar trend as noise increases, although the F1-score exhibits a more stable behavior and does not degrade as sharply as the Challenge score under high noise conditions.

Table 5. Performance of the proposed model and the replicated Zhou model under different noise levels (SNR).

SNR Levels (dB)	**Models**			
	Proposed model		**Replied Zhou Model**	
	Score	**F1-Score**	**Score**	**F1-Score**
Baseline	0.8827	0.9412	0.8395	0.8889
30	0.8827	0.9412	0.8395	0.8889
20	0.8827	0.9412	0.8671	0.8976
15	0.8827	0.9412	0.8395	0.8889
10	0.8827	0.9412	0.8961	0.9062
5	0.8889	0.9492	0.8671	0.8976
0	0.8889	0.9492	0.8671	0.8976
-5	0.8353	0.9310	0.8253	0.8908
-10	0.5856	0.8039	0.7874	0.8908