Monitoring and diagnosis of cardiac anomalies in hospitalized patients using IoT and LSTM neural networks

João Victor Tavares Santos[1], Márcio Guilherme Silvestrini Júnior[1], Lucas Silva de Oliveira[1], Thabatta Moreira Alves de Araujo[2], Eduardo Habib Bechelane Maia[2]

[1] *Department of Mechatronic, Federal Education Center Technology of Minas Gerais, Campus Divinópolis, MG*
R. Álvares de Azevedo, 400 - Bela Vista, Divinópolis - MG
joaojv497@hotmail.com, marcioguilherme712@gmail.com, lqsoliveira@cefetmg.br
[2] *Department of Computing, Federal Education Center Technology of Minas Gerais, Campus Divinópolis, MG*
R. Álvares de Azevedo, 400 - Bela Vista, Divinópolis - MG
thabatta@cefetmg.br, habib@cefetmg.br

Abstract. Continuous monitoring of patients in hospital settings is essential for preventing cardiac disease and ensuring the safety of hospitalized patients. However, current monitoring methods are limited, as they often only monitor patients for short periods, complicating the identification of changes in health status. This article presents a proof of concept for successfully implementing a continuous vital signs monitoring system using the Internet of Things (IoT) coupled with a Long Short-Term Memory (LSTM) neural network for detecting cardiac anomalies. The neural network development relied on a meticulously labeled database comprising 12-lead electrocardiogram (ECG) signals from 45,152 patients, collected at a sampling rate of 500 Hz. These data were labeled by experts, ensuring their high quality and reliability. This database is fundamental for training machine learning models with high diagnostic precision, enabling accurate identification of cardiac rhythm patterns that indicate potential adverse conditions or deviations from a normal ECG pattern. A robust system architecture was necessary for handling and analyzing the data to utilize this extensive data effectively. Microcontrollers played a crucial role in our system, facilitating data transmission to a web server using the Message Queuing Telemetry Transport (MQTT) protocol, an efficient and fast method for message transmission. The MQTT server, acting as a central hub, received the information and distributed it to two destinations: a real-time monitoring site and the neural network model system. In this environment, the model processed the information and promptly identified any irregularities in the data. The results suggest that the model achieved high precision in identifying these conditions, enabling the issuance of alerts and the implementation of preventive measures with a high efficacy rate. Such actions reduce patient health risks by allowing real-time detection and remote communication of vital signs to healthcare professionals.

Keywords: IoT, IA, HealthTech

1 Introduction

Traditional monitoring methods in hospital environments rely on regular visits from healthcare professionals to patients' bedsides. They can be limited in their effectiveness, as patients are only monitored for short periods, making early detection of changes in their health status difficult. Additionally, it is common for many monitoring systems to initiate diagnostic tests only after identifying clinical symptoms, which is problematic in asymptomatic or atypical conditions, increasing the risk of undetected issues. The effectiveness of these methods is also intrinsically linked to the competence and training of healthcare professionals, as inadequate training can result in serious errors or the omission of new cases [1].

Heart disease (HD) represents a serious public health challenge, affecting millions of people globally, as reported by the World Health Organization (WHO) [2]. Within the Medical Internet of Things (MIoT) system, early detection of HD is crucial for effective treatment and patient recovery [3]. Experts and physicians using MIoT systems have developed various non-invasive approaches to classify and diagnose heart diseases [4]. For example, Liu et al. developed SmartVest, an IoT device for early detection of cardiovascular diseases with ECG sensors, Bluetooth and WiFi connectivity, and cloud data processing. The system addresses challenges in ECG signal quality and proposes a hybrid method to classify signals and accurately detect QRS complexes, showing great potential for large-scale cardiovascular monitoring. Zhang et al. developed PSRWT technology, a portable ECG monitoring system with seven leads and WiFi support. This low-power system transmits ECG data to a specialized monitoring center, allowing doctors to remotely monitor patients' health and alert them to suspicious signs, reducing mortality from cardiovascular diseases. On the other hand, Saadatnejad et al. presented an innovative algorithm for ECG classification on wearable devices with limited processing capabilities. Using wavelet transform and LSTM, the algorithm continuously and locally monitors cardiac activity while preserving data privacy. The proposal stands out for its accuracy in arrhythmia classification and computational efficiency, enabling precise and continuous monitoring. Another notable study is by Nurmaini et al., who introduced an innovative method for automatically classifying and delineating electrocardiogram (ECG) signals using bidirectional LSTM recurrent neural networks. The process includes noise cancellation, waveform segmentation, signal classification into four classes (P wave, QRS complex, T wave, and isoelectric line), and model evaluation. The results showed that the method is highly accurate under normal and abnormal conditions, suggesting its potential for cardiac monitoring and clinical diagnosis. The bidirectional LSTM approach significantly improves the detection of P waves, QRS complexes, and T waves, offering more reliable and efficient diagnostic tools for the early detection of cardiac anomalies and improving patient care.

Despite technological advancements in hospital equipment that accurately capture vital signs, many hospitals still need help with remote monitoring and cardiac data analysis [9]. This paper proposes to overcome these limitations by implementing a continuous vital signs monitoring system using the Internet of Things (IoT) and Long Short-Term Memory (LSTM) neural networks. IoT allows continuous real-time health data collection through connected devices, which integrate into the hospital environment. A website receives and displays this data in real-time, enabling remote monitoring. An LSTM network processes this data, which is particularly suitable for detecting events based on temporal data sequences, and allowing early identification of cardiac anomalies.

The rest of the paper is structured as follows: Section 2 introduces the proposed approach architecture and data analysis. Section 3 conducts simulations and analyzes the results. Section 4 concludes this paper and reflects on the next steps.

2 Proposed Approach

This section details the methodology, focusing on the system architecture data collection, and the neural network implementation, including the data preparation model, parameters and training techniques.

2.1 Architecture

The system comprises several fundamental components for scalability, reliability, and efficiency. Microcontrollers are low -power consumption devices that collect real-time data and are suitable for constrained environments. The MQTT server utilizes a lightweight communication protocol for IoT networks, ensuring efficient real-time communication between devices and the server [ZOLETT]. AngularJS is a JavaScript framework that facilitates the creation of interactive and dynamic single-page applications for data visualization. Redis, an inmemory database, is used for data caching, providing quick access to frequently used information. RabbitMQ is a message queue management system ensuring reliable and scalable delivery. The LSTM neural network is specialized in temporal sequences, making it ideal for heart signal analysis. The backend handles business logic and data processing, ensuring efficient interaction between system components. Lastly, PostgreSQL is a robust and flexible relational database that supports large volumes of data and complex operations. Implementing asynchronous processing is crucial to enhance performance and maintain its responsiveness.

Asynchronous processing allows non-blocking, operations, ensuring efficient performance and maintaining responsiveness. This approach ensures tasks are performed in the background, maintaining responsiveness in complex systems. Figure 1 illustrates the proposed architecture, highlighting the scalable interaction between components. This architecture allows the addition of new server instances to meet increasing demand, effectively distributes the workload, and ensures that the system can scale without compromising performance. IoT devices send real-time data via the MQTT server. A specialized consumer manages this data, using Redis for temporary storage, and quick processing. The LSTM neural network receives requests via RabbitMQ, returning analyses processed by the backend, enabling continuous real-time evaluation.

2.2 IoT Hardware Board

The electrocardiograph's serial output facilitates data collection using a microcontroller like the ESP32 board. An interface and a microcontroller, such as the ESP32 board, capture the data. The process involves connecting

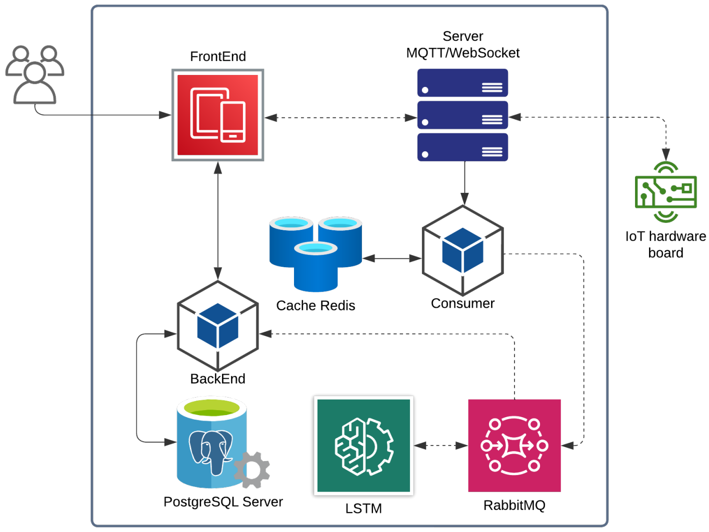

Figure 1. Architecture diagram

the electrocardiograph to a converter that adjusts the voltage levels to be compatible with microcontrollers. The microcontroller is responsible for capturing data from existing hospital equipment and transmitting it via the Internet. Obtaining electrocardiogram readings from humans requires adherence to specific standards and regulations. Therefore, the data embedded in ESP32 was extracted from a pre-existing database [11]. These records were then stored directly on the ESP32, allowing tests to be conducted without involving patients.

The database contains 12-lead electrocardiogram signals collected from 45,152 patients. This resource is important due to its high sampling rate of 500 Hz and the quality of the labels diagnosed and annotated by experts in the field. Data was collected at 32-bit resolution, with amplitudes ranging from -32.768 to 32.767 microvolts, capturing subtle cardiac signal variations. These databases are online and available in [11].

2.3 Data Analysis

We selected tags for normal and abnormal diagnoses during data preprocessing to balance the classes for effective training and validation. Additionally, two tags indicate a normal diagnosis: *426783006* (Sinus Rhythm) and *426177001* (Sinus Bradycardia), while the other tags are associated with abnormal diagnoses. The complete list is available in PhysioNet.

2.4 LSTM neural network

To prepare ECG data, we extracted diagnoses, patient age, and gender to determine record status based on predefined diagnostic codes.

To mitigate class imbalance and *overfitting* issues, we used the *RandomUnderSampler* technique to train the model with balanced data, ensuring an even distribution between normal and abnormal records. This enabled using Long Short-Term Memory (LSTM) Neural Networks for ECG analysis.

LSTM Neural Networks consist of interconnected recurrent units known as LSTM cells. Each cell has a complex structure that enables long-term information memorization and processing. Figure 2 illustrates the LSTM structure, including:

- Forget Gate: Decide which information to retain.
- Input Gate: Controls new information entry.
- Output Gate: Determines cell state contribution to output.
- Connections: Allow communication between cells.

The network architecture comprises an LSTM with three hidden layers, each containing 320 hidden units. The network processes input sequences of dimension 5000, which indicates that each time step in the sequence has 5000 features. After the sequence process, the last output of the LSTM is passed through a fully connected layer that reduces the dimension to two values, suitable for a binary classification task. This configuration captures complex temporal dependencies. The results obtained from the analyzed data will be presented in Section 3.

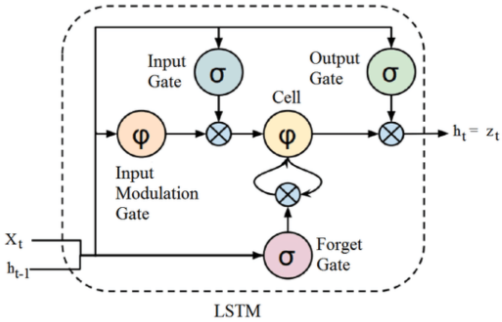

Figure 2. LSTM Network Structure.

Validation and Results

The model parameters were validated by dividing the data into training, validation, and test sets, with 20% of the data belonging to the latter category. The RandomUnderSampler technique was applied to correct the class imbalance, ensuring the sets had a balanced distribution of normal and abnormal diagnoses.

The LSTM model's training was monitored by generating loss and accuracy graphs over the training epochs. The loss curve evaluated the model's error at each epoch, and the accuracy curve monitored the percentage of correctly classified examples. Analyzing these curves was essential to ensure that the model was learning effectively without showing signs of overfitting.

After training the model, the process generated a confusion matrix to quantify the model's performance in terms of correct and incorrect classifications for each class. The confusion matrix allowed a detailed analysis of the model's strengths and weaknesses, providing essential information for calculating performance metrics such as precision, recall, and F1-score.

The Receiver Operating Characteristic (ROC) curve evaluated the binary classification model's overall effectiveness. This curve illustrates the relationship between the false positive rate (FPR) and the true positive rate (TPR) as the decision threshold is adjusted. Calculating the Area Under the ROC Curve (AUROC) metric quantified the model's overall performance.

3 Results

The implementation of the IoT platform proved to be successful and functional. The implementation of the proposed architecture is complete, and all components operate as expected. Testing the platform in a controlled environment showed that all connected devices collected and transmitted data in real time without any failures. It is important to note that the system refrained from directly collecting data to adhere to patient protection laws; instead, a database collects data, and the information is transmitted.

3.1 FrontEnd

The user interface of the monitoring system allows real-time visualization of the collected health data. The figures in the repository [13] include screenshots of the front-end interfaces, such as login screens, alerts, and real-time monitoring. The interface was developed using the AngularJS framework, which allowed for the creation of an interactive and responsive web application. This application facilitates data visualization and navigation through the various monitored parameters. The integration with the MQTT server and the backend occurred efficiently, ensuring continuous updates of the displayed information. Additionally, it informs users when a device is *online*.

3.2 Data Transmission

The electrocardiogram (ECG) signals used in the tests were extracted from a public database due to the unavailability of hospital equipment for direct signal collection. The microcontroller continuously transmits the data to the central system, which processes the information in real-time. Due to the high frequency of the signals, the data was sent in packets, allowing data collection at the hospital equipment's sampling rate and transmission at a lower frequency. This transmission method ensured no significant data loss and enabled rapid transmission.

Thus, the limitation of the lack of direct collection equipment was overcome by using a robust database and efficient transmission to ensure the quality and continuity of the data for analysis.

3.3 LSTM Network

Figure 3 shows a graph developed to validate the parameters of the trained model. This graph demonstrates the progression of the model's loss and accuracy characteristics during the training epochs for both the training and validation datasets.

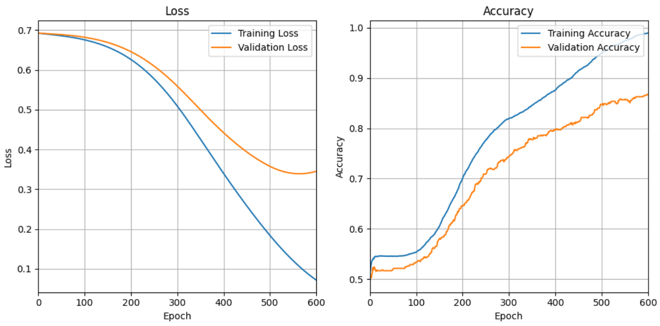

Figure 3. Loss/Accuracy Graph of the trained model.

The loss graph represents the model's error at each epoch. Ideally, the loss curve should decrease monotonically throughout the training, indicating that the model is approaching the optimal solution. The loss curve shows a constant decline, demonstrating that the model is effectively learning from the data and indicating that the model has achieved good generalization for the training data. The accuracy graph represents the percentage of examples the model correctly classifies at each epoch. Ideally, the accuracy curve should increase during the training, indicating that the model is improving its ability to discriminate between classes. Additionally, the model was programmed to employ early stopping to prevent overfitting. This mechanism halted the training process when the validation loss started to rise, ensuring the model did not continue training unnecessarily after reaching its optimal point, thus improving generalization. Analysis of the loss curve for training and validation shows no *overfitting*, indicating good generalization for applied cases.

3.4 LSTM Network Performance

The collected and analyzed data demonstrated that the system accurately identified cardiorespiratory anomalies. The accuracy of the monitoring system was validated through comparisons with conventional diagnostic systems in a test set with data that the model had no previous access to, showing that the IoT and neural network-based approach can be an effective alternative for the early detection of anomalies.

Figure 4 shows a confusion matrix created to quantify the network's performance. This matrix represents the number of correct and incorrect results for each class, allowing the identification of the model's strengths and weaknesses.

Based on the Figure 4, the following performance metrics can be calculated:

$$P = \frac{T_p}{T_p + F_p} = \frac{301}{301 + 38} = 0.887 \tag{1}$$

where P is the precision, T_p are the true positive results, and F_p are the false positive results.

$$R = \frac{T_p}{T_p + F_n} = \frac{301}{301 + 49} = 0.86 \tag{2}$$

where R is the recall, and F_n are the false negative results.

$$F_1 = 2\frac{PR}{P + R} = 2\frac{0.887 \cdot 0.86}{0.887 + 0.86} = 0.873 \tag{3}$$

where F_1 is the F1-score.

Data analysis indicates that the precision of 0.887 means the model is correct 88.7% of the time when classifying an instance as positive. The F1-score of 0.873 is a balanced measure of precision and recall, indicating that the model performs well in classifying both classes.

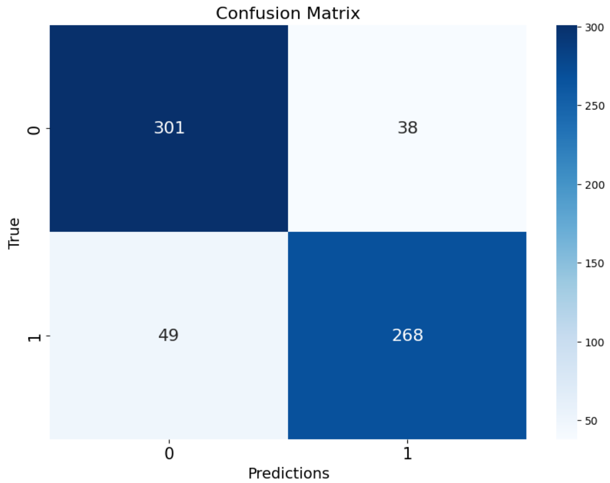

Figure 4. Confusion Matrix of the LSTM network.

The performance analysis of the LSTM neural network, using the ROC (Receiver Operating Characteristic) curve, is fundamental to evaluating the effectiveness of a binary classification model. The ROC curve illustrates the relationship between the false positive rate (FPR) and the true positive rate (TPR) as the decision threshold is adjusted. The AUROC (Area Under the ROC Curve) metric quantifies the overall performance of the model, with values ranging from 0.5 (random performance) to 1.0 (perfect performance). The analysis of the ROC curve of the network, shown in Figure 5, indicates a performance of 0.94, demonstrating a high capacity for discriminating between positive and negative classes. The smoothness and consistency of the ROC curve suggest good model generalization, with no signs of overfitting to the training data. In conclusion, the LSTM neural network presents the desired performance, effectively distinguishing between classes.

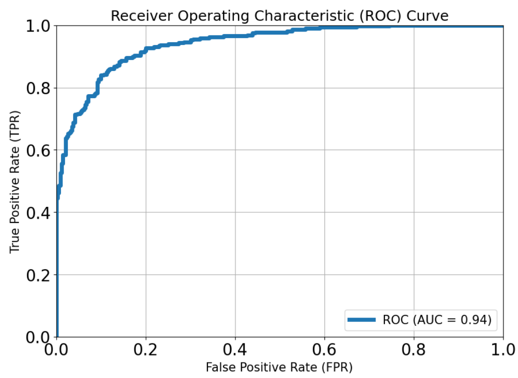

Figure 5. AUROC Curve of the LSTM network.

This study adopted a different approach by separating the dataset into normal and abnormal cases, in contrast to other research that uses specific disease labels from the same dataset. Consequently, there is no direct performance comparison with other works due to the distinct methodology employed. This approach allows the model to broadly identify cardiac anomalies without relying on specific labels. This is particularly advantageous in contexts where greater flexibility and applicability of the model in clinical scenarios are needed.

4 Conclusion

The methodology demonstrates the proposed system's viability for monitoring and diagnosing cardiac anomalies in real-time, contributing significantly to the healthcare area, by utilizing advanced data collection techniques with electrocardiograph equipment, a comprehensive and high-quality dataset was generated, comprising 12-lead electrocardiogram signals from 45,152 patients, sampled at 500 Hz and annotated by field experts. The data preprocessing phase meticulously balanced the classes using the RandomUnderSampler technique, ensuring an even distribution of normal and abnormal records. This balanced dataset was crucial for training the LSTM (Long Short-Term Memory) Neural Networks, which are well-suited for handling time-series data due to their ability to memorize and process long-term information. The validation process involved dividing the dataset into training, validation and test sets and continuously monitoring the model's performance using loss and accuracy curves to avoid overfitting. Generating a confusion matrix using the test set provides crucial insights into the model's strengths and weaknesses, facilitating the calculation of vital performance metrics such as precision, recall, and F1-score. The overall effectiveness of the binary classification model was further validated using the ROC curve and the AUROC metric, which quantified the model's discriminative power. In summary, the comprehensive and systematic methodology ensures that the system can effectively differentiate between normal and abnormal cardiac signals, providing a reliable tool for real-time cardiac anomaly detection and contributing to advancements in healthcare diagnostics. Exploring and comparing different types of neural network architectures is recommended to continue this research. These may include Convolutional Neural Networks (CNNs), Recurrent Convolutional Neural Networks (CRNNs), as well as Autoencoders, such as Variational Autoencoders (VAEs) and Recurrent Autoencoders. By applying and comparing these approaches, we can obtain a more comprehensive assessment and identify the most effective solution for detecting cardiac anomalies in time series. Additionally, continuing the research could involve developing the electronic aspect. At this stage, creating a complementary circuit of protection circuits is crucial to ensure data integrity and system safety.

However, while these methodologies provide a strong foundation for real-time anomaly detection, they must also address the challenges of transitioning from controlled environments to real-world applications. For applications in human subjects, obtaining approval from relevant regulatory bodies is also mandatory, ensuring compliance with ethical standards and safety regulations.

It is currently unknown what specific issues might arise in real-world environments, as these can only be fully understood once the system is applied in such settings. To address potential data reliability issues, it is essential to recognize that data collected directly from the equipment may include various types of noise and interference not present in a pre-existing database. To handle these real-time challenges effectively, the system must implement advanced signal processing techniques, such as digital filtering and noise reduction algorithms. Future work will explore deploying these techniques in real-world scenarios to capture and address any unforeseen problems, thereby validating the system's robustness against potential inconsistencies and environmental variations.

References

[1] Y. J. Choo, G. W. Lee, J. S. Moon, and M. C. Chang. Application of non-contact sensors for health monitoring in hospitals: a narrative review. *Frontiers in Medicine*, vol. 11, 2024.

[2] S. Mendis, P. Puska, B. Norrving, W. H. Organization, and W. H. Federation. *Global Atlas on Cardiovascular Disease Prevention and Control*. World Health Organization, 2011.

[3] P. M. Kumar and U. Devi Gandhi. A novel three-tier internet of things architecture with machine learning algorithm for early detection of heart diseases. *Computers Electrical Engineering*, vol. 65, pp. 222–235, 2018.

[4] D. Mpanya, T. Celik, E. Klug, and H. Ntsinjana. Predicting mortality and hospitalization in heart failure using machine learning: A systematic literature review. *International Journal of Cardiology Heart & Vascular*, vol. 34, pp. 100773, 2021.

[5] C. Liu, X. Zhang, L. Zhao, F. Liu, X. Chen, Y. Yao, and J. Li. Signal quality assessment and lightweight qrs detection for wearable ecg smartvest system. *IEEE Internet of Things Journal*, vol. 6, n. 2, pp. 1363–1374, 2019.

[6] J. Zhang, R. Sun, S. Wang, and J. Zhang. The design of seven-lead electrocardiograph monitoring system based on wi-fi. In *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pp. 1–5, 2018.

[7] S. Saadatnejad, M. Oveisi, and M. Hashemi. Lstm-based ecg classification for continuous monitoring on personal wearable devices. *IEEE Journal of Biomedical and Health Informatics*, vol. 24, n. 2, pp. 515–523, 2020.

[8] S. Nurmaini, A. E. Tondas, A. Darmawahyuni, M. N. Rachmatullah, J. Effendi, F. Firdaus, and B. Tutuko. Electrocardiogram signal classification for automated delineation using bidirectional long short-term memory. *Informatics in Medicine Unlocked*, vol. 22, pp. 100507, 2021.

[9] V. Nangalia, D. R. Prytherch, and G. B. Smith. Health technology assessment review: Remote monitoring of vital signs - current status and future challenges. *Critical Care*, vol. 14, n. 5, pp. 233, 2010.

[ZOLETT] A. R. G. D.; RAMIREZ. ZOLETT. Desenvolvimento de uma interface de monitoração remota para o sistema robótico robix, integrando o protocolo mqtt e oros.

[11] PhysioNet. A large scale 12-lead electrocardiogram database for arrhythmia study. `https://www.physionet.org/content/ecg-arrhythmia/1.0.0/#files-panel`, 2024a.

[12] PhysioNet. Condition names snomed-ct. `https://www.physionet.org/content/ecg-arrhythmia/1.0.0/ConditionNames_SNOMED-CT.csv`, 2024b.

[13] J. V. T. Santos. Photo gallery project. `https://github.com/Joaovictortav/photo-gallery`, 2024.