2D Image-Based Atrial Fibrillation Classification

Felipe M. Dias[1,2], Nelson Samesima[1], Adele Ribeiro[1], Ramon A. Moreno[1],
Carlos A. Pastore[1], Jose E. Krieger[1], Marco A. Gutierrez[1]
[1] Instituto do Coração, Hospital das Clínicas HCFMUSP, Faculdade de Medicina, Universidade de São Paulo
[2] Escola Politécnica, Universidade de São Paulo

Abstract

Atrial fibrillation (AF) is a common arrhythmia (0.5% worldwide prevalence) associated with an increased risk of various cardiovascular disorders, including stroke. Automated routine AF detection by Electrocardiogram (ECG) is based on the analysis of one-dimensional ECG signals and requires dedicated software for each type of device, limiting its wide use, especially with the rapid incorporation of telemedicine into the healthcare system. Here, we implement a machine learning method for AF classification using the region of interest (ROI) corresponding to the long DII lead automatically extracted from DICOM 12-lead ECG images. We observed 94.3%, 98.9%, 99.1%, and 92.2% for sensitivity, specificity, AUC, and F1 score, respectively. These results indicate that the proposed methodology performs similar to one-dimensional ECG signals as input, but does not require a dedicated software facilitating the integration into clinical practice, as ECGs are typically stored in PACS as 2D images.

1. Introduction

Atrial fibrillation (AF) is the most prevalent cardiac arrhythmia worldwide and is associated with an increased risk of stroke, heart failure, and sudden death [1]. Thus, the identification and prioritization of patients with such a condition is highly desirable.

The diagnosis of AF is based on a standard 12-lead electrocardiogram (ECG) where irregular electrical activity is accompanied by the absence of the P wave [2]. This is a time-consuming task that requires a specialized trained professional.

In emergency departments (ED), ECG screening is often employed especially in patients who report symptoms such as chest pain and palpitations [3]. A major challenge is to stratify the ones with life-threatening heart disease that deserve immediate care. In this context, the development of an automatic screening tool to help in the prioritization of patients at higher risk can benefit and optimize this process.

Towards this end, several groups are using both classical machine learning [4, 5] and deep learning approaches [6–8] to detect atrial fibrillation on one-dimensional ECG signals. On the other hand, in hospital settings, ECG exams are often stored as images in the Picture Archiving and Communication System (PACS) using the Digital Imaging and Communications in Medicine (DICOM) format [9, 10]. Similarly, images from point of care ECG devices can be promptly uploaded in the cloud to be analyzed.

In this work, we want to detect atrial fibrillation from ECG scan images using a Convolutional Neural Network (CNN). As the rhythm information is mainly contained in the DII-long lead, we use only the region of interest (ROI) from the original ECG image as an input to the CNN. Demographic information about the patients including age and gender were also fed into the network. We anticipate that such automated process can contribute to improving the prioritization process of high-risk patients in need of specialized attention.

We present in section 2 the development of the methodology and the dataset used to train the neural network. We report the results in section 3 the discussion and closing are presented in sections 4 and 5, respectively.

2. Materials and Methods

2.1. Materials

The data set used consists of 12-lead resting ECGs, obtained retrospectively from the PACS of a tertiary referral hospital from 2017 to 2020 (Hospital das Clínicas - Instituto do Coração). The exams were acquired on the same machine model, MORTARA ELI 250c, in the time and voltage scales of 25 mm / s and 10 mm / mV, respectively. They are stored in PACS as 3320 x 2219 images. We excluded patients under the age of 18, pacemaker users, and unreported exams from the dataset. After applying these exclusion criteria, the dataset contains 79,077 exams from 64,766 different patients.

The exams were analyzed by a cardiologist from the hospital's medical staff, who provides structured diagnos-

tic reports. In total, there are 52 different possible diagnoses, however, in this work, we focus only on atrial fibrillation.

Table 1 provides an overview of the employed dataset, entitled InCor-Db.

Table 1. Summary of the InCor-Db 12-lead ECG dataset.

		AF	NAF	Total
# of reports		7,695	71,382	79,077
Age		68±13	59±17	60±17
Gender	Male	4,310	36,196	40,506
	Female	3,257	33,949	37,206
	Other	128	1,237	1,365

AF: Atrial Fibrillation
NFA: Non-Atrial Fibrillation

2.2. Methods

As mentioned, the ECG exams are often stored as images in DICOM files. The main objective of our approach is to provide a classification system that is easy to integrate into the hospital's workflow. Figure 1 provides a diagram that illustrates how our approach can be integrated into the hospital workflow. In this figure, we show that ECG exams are stored as DICOM files and sent to PACS. In the proposed approach, we intercept this file and use it to directly classify the exam, without modifying the inner workings of the hospital's data flow.

Next, we present the pre-processing steps that were performed on the ECG image, followed by a description of the neural network architecture. We then show the training and testing strategy using the dataset.

2.2.1. Pre-processing

The ECG imaging exam contains several visual cues to aid clinicians in diagnosis, for example, scale grid and colored background. However, since all database images are in the same format, for neural networks these details are not important. The first step is to pre-process the input image. This step includes transforming the image to grayscale and applying a boundary operation to remove the background grid. In addition, to make the signal traces more evident in the image, a dilation operation is also applied to the image. Finally, as mentioned, the rhythm information is mainly contained in the long DII lead of the ECG, so this region is cut out of the image. To reduce the computational complexity of the model, this region is also reduced to 30% of its original size.

2.2.2. Neural Network

Our proposed classification method consists of a CNN that receives as input the image of the DII lead along with demographic data (age and gender). The network architecture consists of five blocks, each with the following layers: Conv2D, Batch Normalization, ReLU, Conv2D, Batch Normalization, ReLU, and Max Pooling. The fully connected layer receives the output of the last block and demographic data. Figure 2 shows the proposed neural network architecture.

2.2.3. Training and Testing

The Incor-Db dataset was divided into training, validation, and test using 60%, 20%, and 20%, respectively. To prevent data leakage, we ensure that exams from the same patient cannot exist in different divisions. We use Adam's optimizer with a learning rate of 10^{-4} to minimize the binary cross-entropy loss function. In addition, we trained the network for 100 epochs using a 64 batch size.

3. Results

3.1. Experimental Setup

Our experiments were performed using a Foxconn High-Performance Computer (HPC) M100-NHI with an 8 GPU cluster of 16GB NVIDIA Tesla V100 cards. The methodology was implemented using the Python framework and Keras / TensorFlow. To evaluate our approach, we used Sensitivity (Se), Specificity (Spe), Positive Predictive Value (PPV), F1 score (F1), Area Under Operational Receipt Curve (AUC), and Accuracy (Acc).

3.2. Experimental Results

We present a summary of the calculated metrics and the corresponding confusion matrix for the test split of the InCor-Db dataset in Table 2 and Figure 3, respectively.

Table 2. Summary of the obtained results for the InCor-Db test split.

Metric	Value
Sensitivity	94.3%
Specificity	98.9%
Positive Predictive Value	90.2%
F1-score	92.2%
AUC	99.1%
Accuracy	98.5%

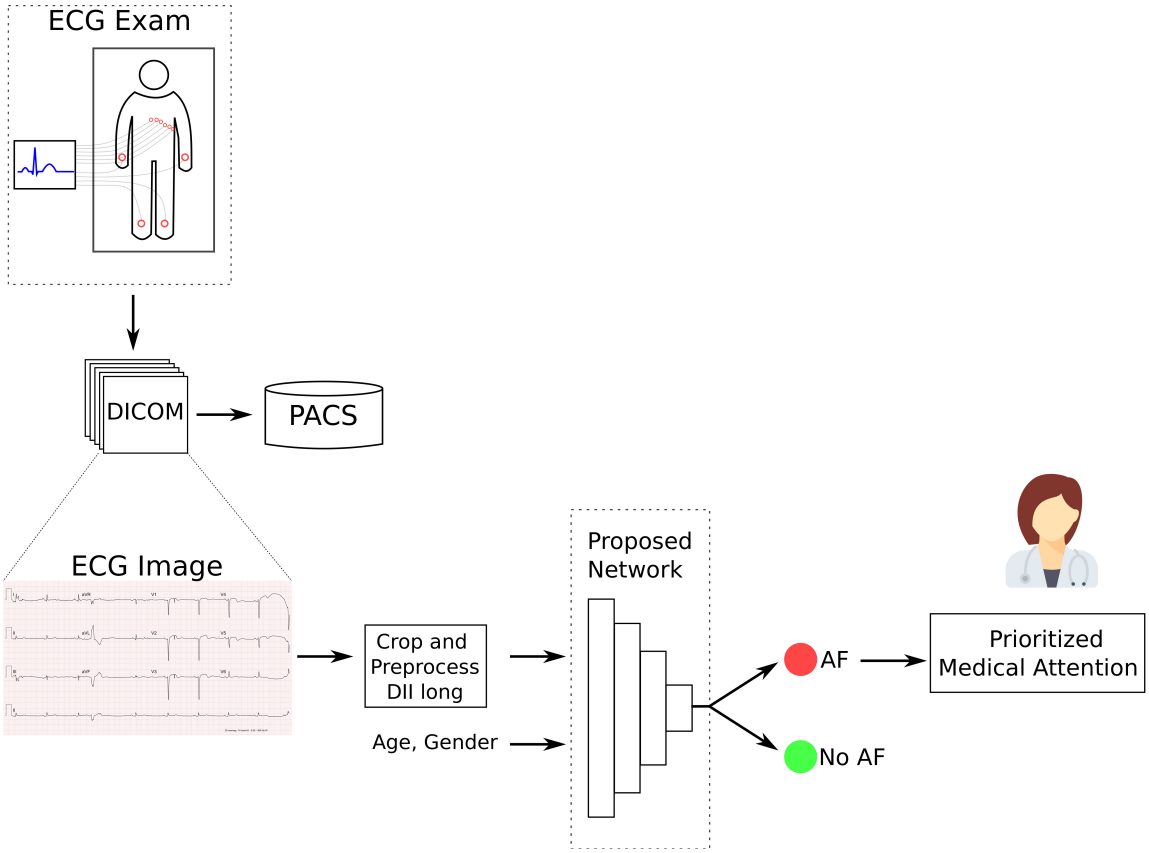

Figure 1. Integration of our proposed approach in a hospital environment.

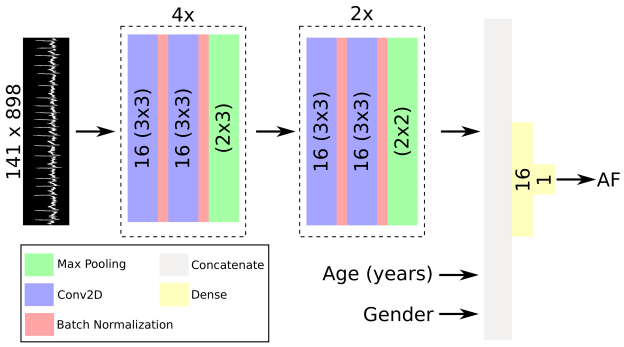

Figure 2. Proposed neural network architecture.

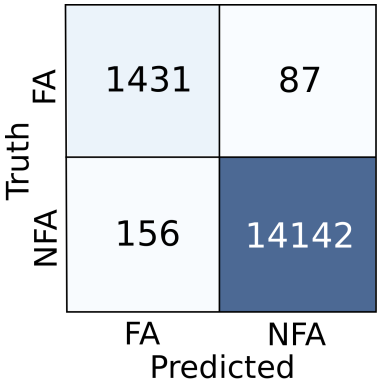

Figure 3. Confusion matrix for the InCor-Db test split.

4. Discussion

Considering that our main goal is to develop a screening tool, a low number of false negatives or a high sensitivity metric indicates that our method is performing well. As shown in Table 2, we obtained results for all metrics considered greater than 90%. In particular, in Figure 3, we show that among 1,518 ECG exams labeled as FA, we found only 87 false negatives, resulting in 94.3% sensitivity.

There are previous reports [8] that obtained a sensitivity of 76.9% for classifying atrial fibrillation from 12-lead ECG signals. Thus, although we did not use the same data set and methodology, we may consider that the reported performance indicates progress. The advantage of the proposed method is its easy integration with the hospital environment, allowing for quick clinical applicability. Furthermore, as the method developed is based on super-

vised learning, this integration would enable a continuous improvement of the neural network as new diagnoses are added by doctors.

One limitation of the present work is that we limited the screening tool to identify a single entity. We chose atrial fibrillation because it is an important and prevalent risk factor and patients with this condition require prompt care. The time for medical care directly impacts the patient's chance of survival in this type of occurrence [11]. Thus, in the future, we want to expand the number of classification classes with a greater focus on the more severe conditions till our methodology can be fully trained to identify the different classes of rhythm abnormalities. We also acknowledge that validation of the methodology with an external dataset is necessary prior to the general application of this methodology.

5. Conclusion

Here we describe an AF classification system from images from ECG exams using CNN. We first select the region corresponding to the long DII lead of the ECG image, as this lead provides most of the rhythm information. This image is then resized and preprocessed to eliminate useless diagnostic information. Pre-processed images were used to train a CNN using an ECG dataset containing nearly 80,000 exams.

Finally, we provide evidence that using this approach we obtained a low number of false negatives, consistent with screening purposes. The use of images as input does not require dedicated software and can be easily integrated into hospital data systems. As the system is used, we anticipated that the performance of the CNN must improve and a similar approach will be used to increase the number of classification classes detected by the CNN.

Acknowledgments

This work was supported by Foxconn Brazil and Zerbini Foundation as part of the research project "Machine Learning in Cardiovascular Medicine".

References

[1] Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. International Journal of Stroke 2021; 16(2):217–221.

[2] Gutierrez C, Blanchard DG. Atrial fibrillation: diagnosis and treatment. American family physician 2011;83(1):61–68.

[3] Su HY, Tsai JL, Hsu YC, Lee KH, Chang CS, Sun CK, Wang YH, Chi SC, Hsu CW. A modified cardiac triage strategy reduces door to ecg time in patients with st elevation myocardial infarction. Scientific Reports 2021; 11(1):1–9.

[4] Christov I, Krasteva V, Simova I, Neycheva T, Schmid R. Multi-parametric analysis for atrial fibrillation classification in ecg. In 2017 Computing in Cardiology (CinC). IEEE, 2017; 1–4.

[5] Coppola EE, Gyawali PK, Vanjara N, Giaime D, Wang L. Atrial fibrillation classification from a short single lead ecg recording using hierarchical classifier. In 2017 Computing in Cardiology (CinC). IEEE, 2017; 1–4.

[6] Dang H, Sun M, Zhang G, Qi X, Zhou X, Chang Q. A novel deep arrhythmia-diagnosis network for atrial fibrillation classification using electrocardiogram signals. IEEE Access 2019;7:75577–75590.

[7] Pourbabaee B, Roshtkhari MJ, Khorasani K. Deep convolutional neural networks and learning ecg features for screening paroxysmal atrial fibrillation patients. IEEE Transactions on Systems Man and Cybernetics Systems 2018; 48(12):2095–2104.

[8] Ribeiro AH, Ribeiro MH, Paixão GM, Oliveira DM, Gomes PR, Canazart JA, Ferreira MP, Andersson CR, Macfarlane PW, Meira Jr W, et al. Automatic diagnosis of the 12-lead ecg using a deep neural network. Nature communications 2020;11(1):1–9.

[9] Sassi R, Bond RR, Cairns A, Finlay DD, Guldenring D, Libretti G, Isola L, Vaglio M, Poeta R, Campana M, et al. Pdf–ecg in clinical practice: A model for long–term preservation of digital 12–lead ecg data. Journal of electrocardiology 2017;50(6):776–780.

[10] Badilini F, Young B, Brown B, Vaglio M. Archiving and exchange of digital ecgs: A review of existing data formats. Journal of electrocardiology 2018;51(6):S113–S115.

[11] Berger PB, Ellis SG, Holmes Jr DR, Granger CB, Criger DA, Betriu A, Topol EJ, Califf RM. Relationship between delay in performing direct coronary angioplasty and early clinical outcome in patients with acute myocardial infarction: results from the global use of strategies to open occluded arteries in acute coronary syndromes (gusto-iib) trial. Circulation 1999;100(1):14–20.

Address for correspondence:

Felipe Meneguitti Dias
Heart Institute (InCor) - Av. Dr. Enéas Carvalho de Aguiar, 44 - Cerqueira César, São Paulo - SP, Brazil
f.dias@incor.usp.br