ImECGnet: Cardiovascular Disease Classification from Image-Based ECG Data Using a Multi-branch Convolutional Neural Network

Amir Ghahremani and Christoph Lofi*

Delft University of Technology, Delft, the Netherlands; Email: a.ghahremani@tudelft.nl (A.G.)
*Correspondence: c.lofi@tudelft.nl (C.L.)

***Abstract*—Reliable Cardiovascular Disease (CVD) classification performed by a smart system can assist medical doctors in recognizing heart illnesses in patients more efficiently and effectively. Electrocardiogram (ECG) signals are an important diagnostic tool as they are already available early in the patients' health diagnosis process and contain valuable indicators for various CVDs. Most ECG processing methods represent ECG data as a time series, often as a matrix with each row containing the measurements of a sensor lead; and/or the transforms of such time series like wavelet power spectrums. While methods processing such time-series data have been shown to work well in benchmarks, they are still highly dependent on factors like input noise and sequence length, and cannot always correlate lead data from different sensors well. In this paper, we propose to represent ECG signals incorporating all lead data plotted as a single image, an approach not yet explored by literature. We will show that such an image representation combined with our newly proposed convolutional neural network specifically designed for CVD classification can overcome the aforementioned shortcomings. The proposed (Convolutional Neural Network) CNN is designed to extract features representing both the proportional relationships of different leads to each other and the characteristics of each lead separately. Empirical validation on the publicly available PTB, MIT-BIH, and St.-Petersburg benchmark databases shows that the proposed method outperforms time series-based state-of-the-art approaches, yielding classification accuracy of 97.91%, 99.62%, and 98.70%, respectively.**

***Keywords*—Convolutional Neural Network (CNN), classification, Electrocardiogram (ECG)**

I. Introduction

A typical ECG signal consists of several simultaneously recorded sensor signals, also known as ECG leads (e.g., 12 or 15 parallel leads is common). Conventional state-of-the-art methods for automatic ECG classification [1–11] represent the ECG signals as time-series of sampled measurements, and design methods to process these data and perform the classification. In [11], authors have developed a multi-scale CNN to process the matrix of input time-series signal and classify the 8-lead ECGs. The work in [12] applies an attention mechanism (to estimate which parts of the time-series signal are more important to focus on during learning) and LSTM to the ECG signal and its wavelet power spectrum. Then, it combines the results to generate the classification. Tan *et al.* [8] proposed a combined implementation of a LSTM network together with CNN for CVD identification from ECG signals. This work uses wavelet transformation to reduce noise from input time-series data. Hasan and Bhattacharjee [9] proposes a method to classify multiple CVDs using a one-dimensional CNN where a modified ECG signal is given as an input signal to the network. The work enhances the input signal quality by reducing the noise using the empirical mode decomposition technique.

Generally, such methods are prone to noise in the recorded signal [9]. To diminish the effect of noise, several works have attempted to employ complementary noise reduction approaches to enhance the signal-to-noise ratio. We propose an alternative approach plotting ECG data points on frames and treat them as graphical images. With this, the dependency of method performance to the input noise is highly decreased because the effect of the noise signal on the whole image is neglectable. Then, we design an elaborated deep-learning network for processing this data representation. According to the empirical validation on three benchmark databases, the proposed method performs better classification compared to the conventional time-series based methods, since this method is more resilient to noise and effects of data pre-processing, and can pick-up significant patterns more effectively compared to operating on raw time-series data.

This paper is organized as follows: Section II explains the proposed method in detail. Section III presents the experimental results and validation.

II. ImECGnet: Image-Based ECG Network

This section presents the ImECGnet. First, we provide an overview of the proposed method. Then, we explain each part of the architecture in detail.

Manuscript received July 1, 2022; revised August 1, 2022; accepted October 9, 2022.

doi: 10.18178/joig.11.1.9-14

A. Architecture Overview

Fig. 1 illustrates the architecture of the proposed ImECGnet. The proposed network receives a mini-batch of generated ECG images as input. Then, it extracts a feature embedding for the input images using a backbone network. Thereupon, the extracted embedding is further investigated using two parallel branches. These branches are elaborately designed to generate more indicative feature representation of input images by further discriminating the extracted feature embedding. The branches process the backbone network's extracted feature embedding to generate specific features of lead-oriented and channel information. This multi-branch type of architecture is inspired by recent methods developed for computer vision re-identification applications like MGN [13] and PRN [14], which show significant improvement in re-identification performance for pedestrian and vehicles, respectively. The proposed ImECGnet exploits two branches and uses a SoftMax loss function to calculate the gradients required in the training.

B. ImECGnet Description

This subsection explains each part of the proposed ImECGnet in detail, as illustrated in Fig. 1.

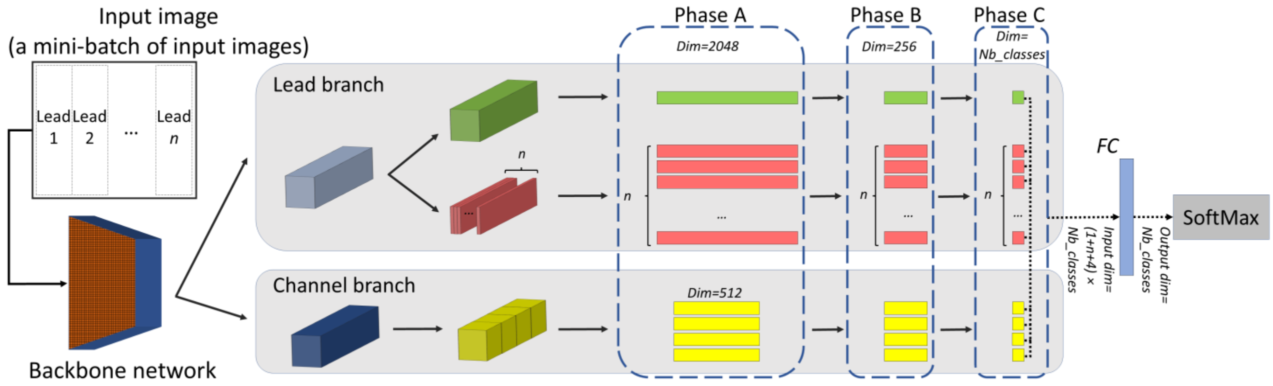

Figure 1. ImECGnet architecture. Lead branch extracts the overall and lead-based features. The channel branch extracts the correlations between feature maps of the CNN. FC stands for the fully connected layer. The parameter *n* represents the number of recorded leads in the original ECG signal (e.g. *n*=15 for the PTB database).

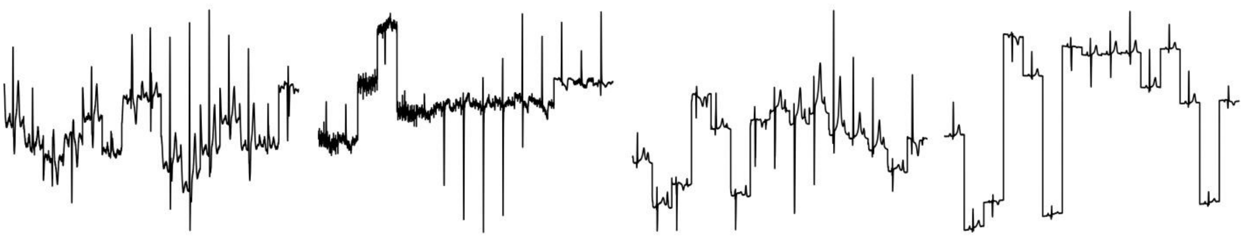

Figure 2. Four image-based ECG examples generated from different ECG signals of the PTB database. Each image is generated by firstly concatenating 1,000 consecutive datapoints of 15 simultaneous leads of the selected ECG signal (i.e. 15,000 datapoints in total), and then plotting them on the same frame. Proportional relationships of leads to each other are sensible in each image (based on the corresponding CVD).

1) Generation of input images

The data in an ECG database are time-series data points, digitized at a specific frequency from recorded ECG signals. As mentioned in previous sections, state-of-the-art methods employ these data as digitized data points (matrix format). In this work, we propose to employ the ECG data as images. To this end, we select 1,000 digitized data points of all ECG leads and concatenate them in a sequential manner (on the horizontal axis). The selected 1,000 digitized data points from leads are chosen from overlapping parts of the simultaneously recorded ECG leads. However, in order to preserve inter/intra-sequence features, these 1,000 data points are chosen from random parts of the input ECG signal. With this, we have applied a data augmentation technique, which also reduces the dependency of the system on specific parts of the ECG signal. This overlapping strategy is adopted from the work in [9].

After generating the concatenated ECG sequences, we add a second dimension for each data point of the sequence, representing the number of the data point in the sequence. This means each data point will now be represented by an ordered pair of (x, y), where x is the number of each data point in the generated ECG sequence and y is the value of the corresponding data point. Then, we plot these ordered pairs on a 2D image frame, where the horizontal and vertical axes represent the x and y dimensions of ordered pairs, respectively. We perform this plotting operation using the plot function of Python, which also applies interpolation between discrete data points. The plotting occurs on image frames of the same size with 600 dpi resolution for all ECG samples. At the end, the generated images have dimensions of 2,199 × 2,725 pixels (height × width). These images will be later downsampled by ImECGnet to a size of 314 × 384 pixels, keeping the aspect ratio. Then, ImECGnet applies conventional standardization techniques on the resized input images. Fig. 2, illustrates few examples of the generated images for different CVD classes of PTB database. Here, it is important to mention that there are few ECG classification works in the literature, which also represent their input ECG data as images [15–17]. However, these works obtain their input images using the spectrogram of ECGs or the graphical print-outs of ECG measuring machines, while our work preserves and plots every ECG data-point in the generated ECG image.

2) ImECGnet architecture

The proposed ImECGnet investigates the input images to extract three types of features: a) overall features, which yield information produced by concatenation of leads altogether, and consequently, represent the proportional relationships of leads to each other, b) lead-based features, which represent the information produced by each ECG lead separately, and c) channel information, which express the correlation existing among feature maps of the CNN. From a design perspective, generally, CNNs are constructed as a sequence of feature extraction layers (e.g. convolutional layers), such that the earlier layers extract shallower features from input images, and then, the following layers extract deeper features upon the shallower features. For example, in a CNN designed for the classification of objects, shallower layers extract common features for all classes (e.g., curves and lines), while deeper layers extract class-specific features (e.g. tires of a car). The mentioned three types of features targeted by the ImECGnet are deep information. Therefore, the design needs a backbone network to extract the shallower features embedding (the coarse portion of the required feature representation). This will be followed by ECG-oriented layers, designed specifically to further process the obtained feature embedding for the intended application. Hence, we explain the designed CNN as a backbone network and a multi-branch network as follows.

a) Backbone network

In order to facilitate the procedure of designing the multi-branch network, an initial set of layers from reliable classification CNNs are employed as a backbone network. This work employs the initial set of layers from the ResNet [18] as the backbone network, which is a well-known network for its robust performance against the vanishing gradients problem. The network has several versions, depending on the number of internal layers like Renet18, ResNet34, ResNet50, etc. For example, in ResNet50, the number 50 represents the number of layers of the network. In order to select the right version for our purpose, we have separately trained ResNet18, ResNet34, and ResNet50 (without the multi-branch design) on our image-based ECG data. After experimental validation, we have noticed that ResNet50 is working best among others on our ECG image data. Therefore, we employ ResNet50 as our backbone network to extract the feature maps for each image of the mini-batch. Then, the extracted feature maps at the end of the Conv4_1 layer (see [18] for details) are duplicated two times and are fed separately into the two branches. In other words, these branches will be developed on top of separate copies of the Conv4_1 layer of the ResNet50. In the remainder of this paper, we will refer to each of these duplications of feature maps as a feature embedding. Both the lead and channel branches will then further process their embeddings to construct meaningful features in a parallel manner.

b) Multi-branch design

The core contribution of ImECGnet is its proposed 2-branch architecture, specifically designed to address the ECG classification problem. These branches are named as a lead branch and channel branch (as illustrated in Fig. 1), which concentrate on the lead and channel-wise structures existing in their input feature embedding, respectively.

The lead branch performs two independent partitioning operations on the width dimension of its input feature embedding. The outcomes of these operations are 1 and n separate feature volumes. Here, the parameter n represents the number of leads per each ECG signal in the original database. The n value for PTB, MIT-BIH, and St.-Petersburg benchmark databases is 15, 2, and 12, respectively. The partitioning operation into n separate feature volumes generates volumes that contain an equal fraction of the input feature embedding. In the lead branch, partitioning the input feature embedding into 1 volume aims in duplicating the whole input feature embedding into a separate volume, which intends in preserving the global features of the embedding for the following steps. In the remainder of this paper, we will call this duplicated volume (the 3D block depicted at the leftmost side of the branch) the global volume. The motivation behind partitioning the global volume into n horizontally-partitioned volumes is the following. The n ECG leads are concatenated on the horizontal axis while generating ECG images. Therefore, this horizontal partitioning will separately give the chance to each lead to play a role in the final decision. In a parallel manner, the channel branch performs the operation of partitioning its input feature embedding into 4 equally-sized volumes of features. This partitioning occurs across the depth dimension of the input feature embedding, where the different feature vectors of the CNN are stacked. This partitioning intends at discovering the correlations between different feature maps of the feature embedding.

As illustrated in Fig. 1, generating these n+5 partitioned volumes out of the input feature embedding is followed by a three-phase procedure of further processing the resulting feature vectors. Phase A, separately shrinks the generated feature volumes into 1D feature vectors using global max-pooling operations. The resulting feature vectors have the size of 2,048 and 512 for the lead and channel branches, respectively. Then, phase B equalizes the size of the constructed feature vectors to a size of 256 using 1×1 convolutional operations. This phase also applies a batch normalization operation on the resulting feature vectors. Afterwards, phase C applies a separate fully-connected (FC) layer to the feature vectors constructed in phase B. The output dimension of these FC layers is equal to the number of ECG classes (*Nb_classes*). The outputs of the FC layers provide the score of each feature vector for each class. These outputs are processed by a separate FC layer, depicted in at the rightmost side of the diagram in Fig. 1, before SoftMax block. The input and out dimensions of this FC layer are (n+5) × *Nb_classes* and *Nb_classes*, respectively. The parameter *Nb_classes* stand for the number of CVD classes for the database. In training time, these outputs of this FC layer are transferred into a SoftMax CE loss block to generate the gradients for training the network. In test time, these outputs are employed in making the final classification decision.

III. EMPRICAL VALIDATION

This section presents the experimental material and results. First, the employed databases are introduced. Then, the evaluation settings are explained. Finally, the empirical validation of the proposed method is presented.

A. Benchmark Databases

This work evaluates the proposed ImECGnet method on three publicly available benchmark databases. The first database is the PTB [19] database, which encompasses 549 ECG records from 290 subjects. Each record contains 15 simultaneously measured signals (i.e. ECG leads). The database contains 9 classes of CVDs. The second database is MIT-BIH [20] database. This database contains 48 half-hour ECG measurements with two leads, digitized at 360 samples per channel from 47 subjects. This database includes 4 classes of CVDs. The third database is St.-Petersburg database [21], which contains 75 half-hour 12-leads ECGs, digitized at 257 Hz. This database is comprised of 9 CVD classes. The class labels for some of the ECG recordings of the MIT-BIH and St.-Petersburg databases are not provided in the original database.

B. Evaluation Setup

To the best of our knowledge, the work in Hasan and Bhattacharjee's [9] provides the best results for CVD classification from ECG signals among state-of-the-art methods, which is evaluated on the PTB, MIT-BIH, and St.Petersburg databases. We will refer to that work as Ref-ECG in the following, and use it as an evaluation baseline. We mimic the evaluation setup in Ref-ECG for easier comparison.

Ref-ECG applies a specific CNN to ECG samples generated in a 1D vector format with a length of 1,000 data points extracted from original ECG signals. To generate the 1D input vectors, the mentioned work has performed an element-wise summation over simultaneously recorded data points from all leads of an ECG signal. Here, it is important to mention that the authors of the work in Hasan and Bhattacharjee's [9] also introduce a specific transform which reduces the noise level of the raw ECG signals. Evidently, the Ref-ECG can yield a better performance using this noise reduction technique. However, we only train and test the Ref-ECG method on ECG samples generated from raw ECG recordings, as is the case for ImECGnet. With this, we aim to prove the efficiency of image representation-based ImECGnet in dealing with noisy ECG signals.

However, the results originally reported by Ref-ECG in Hasan and Bhattacharjee's [9] are not fully reproducible: 1) The methodology of creating training data is not described in detail. Some records of MIT-BIH and St.Peterburg datasets are unlabeled, and there is no information on how these records were treated. Also the overlap values used while extracting the 1,000 data points of each ECG sample from the longer original ECG signals were not provided. Furthermore, two classes of the PTB database are omitted without further explanation. Thus, we could not recreate the same training and test samples using the graphical feature representation used by ImECGnet; 2) The provided code base is outdated and could not be executed without failure.

Therefore, we opted for method reproducibility of Ref-ECG, creating new test and training datasets from the three benchmark databases, and using our own re-implementation of Ref-ECG as a baseline for our experiments. We have asked a medical doctor to provide the CVD labels for the non-labeled records of the MIT-BIH and the St.-Petersburg database[1]. Then, we generated both the 1D vector-based (for Ref-ECG) and image-based (for ImECGnet) training and test sets with new overlap values for both methods. In order to have a fair comparison, we aim to generate training and test sets which include the exact same ECG samples (i.e. ECG samples extracted from the same parts of original lengthy ECG recordings). All of these ECG samples have the length of 1,000 data points. The only difference between the samples of the training and test sets provided for Ref-ECG and ImECGnet is that the samples generated for Ref-ECG are in 1D vector format (1 × 1,000 vector created by element-wise summation of the corresponding ECG leads, as in [9]), while the samples generated for ImECGnet are in 2D image format, as explained in Section II. At the end, we have trained and tested both the Ref-ECG and ImECGnet methods by implementing and running the required codes on our machine.

In addition to the Ref-ECG, we have also developed two naive image-based baseline models to showcase the benefits of the aforementioned ImECGnet optimizations. These baselines, which are ResNet34 and ResNet50, vary in terms of complexity and having a different number of network layers. These networks (which are pre-trained on ImageNet) are re-trained on our generated image-based ECG datasets. The difference between ImECGnet and these baseline networks is the unique multi-branch architecture of the ImECGnet.

C. Validation Results and Discussion

In order to have a fair comparison between the proposed ImECGnet and other employed methods, we have employed the same training settings. For the training purpose of all these networks, we have employed the stochastic gradient descent optimizer with momentum for 12 epochs. The initial learning rate is set to 1×10^{-3}, which is reduced to one-tenth of this value after every 3 epochs.

Table I presents the CVD classification results. The Ref-ECG is adopted from literature, while rest of the methods in the table are extracted from our research work. According to the table, all the methods based on image

[1] These labels are generated by a medical doctor, as mentioned in the acknowledgment.

representation of ECG samples outperform the time-series based Ref-ECG. For example, the ImResNet50 (i.e., baseline ResNet50 applied to image ECG data) generates approximately 18% higher accuracy compared to Ref-ECG on the St.-Petersburg database. Additionally, ImECGnet performs better than all methods, including image-based baseline ImResNet34 and ImResNet50. For example, ImECGnet provides almost 3% higher accuracy on the St.-Petersburg database compared to ImResNet50. This shows the following two aspects: a) the advantages of treating ECG leads as visual images by classification system, and b) the capacity of multi-branch ImECGnet in processing the image-based ECG data using overall, lead-based, and channel-based features.

TABLE I. IMECGNET PERFORMANCE ANALYSIS. THE PROPOSED METHOD PROVIDES HIGHER ACCURACY COMPARED TO OTHER METHODS TRAINED ON THE SAME DATA AND UNDER THE SAME SETTINGS

Metric	Accuracy (%)		
Database	PTB	MIT-BIH	St.-Petersburg
Ref-ECG [9]	88.04	97.95	77.62
ImResNet34	95.46	99.40	93.49
ImResNet50	96.83	99.37	95.76
ImECGnet	97.91	99.62	98.70

This paper proposes an automated CVD classification of ECG recordings. In contrast to established work, the proposed method represents input ECG records as images, which are generated by concatenating the data points of different leads of ECG recordings and plotting them on image frames. To the best of our knowledge, this paper is the first reported research work that represents ECG records by concatenating the lead data and plotting them on images. Second, the proposed ImECGnet is specifically designed to process such kind of the data. This network is a CNN which further investigates the extracted feature embedding to obtain information representing a) the correlation of different leads with each other, b) each lead's characteristics, and c) channel-oriented properties. The proposed method outperforms state-of-the-art methods by producing up to 21% higher classification accuracy. This proves both the potential of the generated image-based representation in ECG analysis and the efficiency of the proposed ImECGnet. It is also important to mention that our approach is able to perform classification based on very short input signals (e.g. 1 second) since it does not rely on a long-time sequence of data points. For future work, the design of the proposed network can be more elaborated to provide even higher performance.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

A.G. conducted the research, and wrote the paper. C.L. reviewed the paper. All authors had approved the final version.

ACKNOWLEDGMENT

The authors wish to thank Dr. Sina Sadeghimehr for his help in providing the right labels for the samples with no labels.

REFERENCES

[1] A. A. Bhurane, M. Sharma, S.-T. Ru, and U. R. Acharya, "An efficient detection of congestive heart failure using frequency localized filter banks for the diagnosis with ECG signals." *Cognitive Systems Research*, vol. 55, pp. 82–94, 2019.

[2] J. Liu, C. Zhang, Y.-J. Zhu, T. Ristaniemi, T. Parviainen, and F.-Y. Cong, "Automated detection and localization system of myocardial infarction in single-beat ECG using dual-Q TQWT and wavelet packet tensor decomposition," *Computer Methods and Programs in Biomedicine*, vol. 184, 105120, 2020.

[3] Z. I. Attia, S. Kapa, and F. Lopez-Jimenez, P. M. McKie, D. J. Ladewig, G. Satam, P. A. Pellikka, *et al.*, "Screening for cardiac contractile dysfunction using an artificial intelligence — Enabled electrocardiogram," *Nature Medicine*, vol. 25, no. 1, pp. 70–74, 2019.

[4] M. Sharma and U. Rajendra Acharya, "A new method to identify coronary artery disease with ECG signals and time — Frequency concentrated antisymmetric biorthogonal wavelet filter bank," *Pattern Recognition Letters*, vol. 125, pp. 235–240, 2019.

[5] Z. Golrizkhatami and A. Acan, "ECG classification using three-level fusion of different feature descriptors," *Expert Systems with Applications*, vol. 114, pp. 54–64, 2018.

[6] U. R. Acharya, H. Fujita, S. L. Oh, U. Raghavendra, J. H. Tan, M. Adam, A. Gertych, and Y. Hagiwara, "Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network," *Future Generation Computer Systems*, vol. 79, pp. 952–959, 2018.

[7] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, and R. S. Tan, "Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals," *Applied Intelligence*, vol. 49, no. 1, pp. 16–27, 2019.

[8] J. H. Tan, Y. Hagiwara, W. Pang, I. Lim, S. L. Oh, M. Adam, R. S. Tan, M. Chen, and U. R. Acharya, "Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals," *Computers in Biology and Medicine*, vol. 94, pp. 19–26, 2018.

[9] N. I. Hasan and A. Bhattacharjee, "Deep learning approach to cardiovascular disease classification employing modified ECG signal from empirical mode decomposition," *Biomedical Signal Processing and Control*, vol. 52, pp. 128–140, 2019.

[10] B. Murugesan, V. Ravichandran, K. Ram, S. P. Preejith, J. Joseph, S. M. Shankaranarayana, and M. Sivaprakasam, "Ecgnet: Deep network for arrhythmia classification," in *Proc. 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, IEEE, 2018, pp. 1–6.

[11] H. Zhang, W. Zhao, and S. Liu, "SE-ECGNet: A multi-scale deep residual network with squeeze-and-excitation module for ECG signal classification," in *Proc. 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, December 2020, pp. 2685–2691.

[12] S. Mousavi, F. Afghah, A. Razi, and U. R. Acharya, "ECGNET: Learning where to attend for detection of atrial fibrillation with deep visual attention," in *Proc. 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, IEEE, 2019, pp. 1–4.

[13] G. S. Wang, Y. F. Yuan, X. Chen, J. W. Li, and X. Zhou, "Learning discriminative features with multiple granularities for person re-identification," in *Proc. the 26th ACM International Conference on Multimedia*, 2018, pp. 274–282.

[14] H. Chen, B. Lagadec, and F. Bremond, "Partition and reunion: A two-branch neural network for vehicle re-identification," in *CVPR Workshops*, 2019, pp. 184–192.

[15] V. Gupta, M. Mittal, V. Mittal, and A. U. Gupta, "ECG signal analysis using CWT, spectrogram and autoregressive technique." *Iran Journal of Computer Science*, vol. 4, no. 4, pp. 265–280, 2021.

[16] Y. Arpitha, G. L. Madhumathi, and N. Balaji, "Spectrogram analysis of ECG signal and classification efficiency using MFCC

feature extraction technique," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 2, pp. 757–767, 2022.
[17] V. Sangha, B. J. Mortazavi, A. D. Haimovich, A. H. Ribeiro, C. A. Brandt, D. L. Jacoby, W. L. Schulz, H. M. Krumholz, A. L. P. Ribeiro, and R. Khera, "Automated multilabel diagnosis on electrocardiographic images and signals," *Nature Communications*, vol. 13, no. 1, pp. 1–12, 2022.
[18] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778. 2016.
[19] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet," *From the Journal*, pp. 317–318, 1995.
[20] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.
[21] T. Viktor and A. Khaustov, "St.-Petersburg institute of cardiological technics 12 — lead arrhythmia database," *Circulation-Electronic*, vol. 101, no. i23, pp. e215–e220, 2000.

Amir Ghahremani received his MSc. degree in electrical engineering at the Khajeh Nasir Toosi University of Technology (KNTU), Tehran, Iran in 2014. He finalized his MSc. thesis project on image processing and computer vision topics. From his thesis, he extracted four IEEE conference papers on video coding and vehicle tracking, including an IEEE ICIP paper on motion estimation for video coding. This was followed by a PhD in the Video Coding and Architectures (VCA) research group at the Eindhoven University of Technology (TU/e), Eindhoven, the Netherlands. His PhD research interests included computer vision, machine learning, and deep learning. In this period, he published 11 scientific papers in international journals and conferences. Since 2020, he has been working in WIS (Web Information Systems) group of the Delft University of Technology, Delft, the Netherlands, as a postdoctoral fellow. The research field of his postdoc includes medical image and signal processing using deep learning techniques.