

Enhancing Deep Learning-based 3-lead ECG Classification with Heartbeat Counting and Demographic Data Integration

Khiem H. Le ^{a,b,*}, Hieu H. Pham ^{a,b}, Thao BT. Nguyen ^{a,b}, Tu A. Nguyen ^{a,b}, Tien N. Thanh ^c, Cuong D. Do ^{a,b}

^aVinUni-Illinois Smart Health Center, VinUniversity, Hanoi, Vietnam

^bCollege of Engineering and Computer Science, VinUniversity, Hanoi, Vietnam

^cCollege of Health Sciences, VinUniversity, Hanoi, Vietnam

Abstract—An increasing number of people are being diagnosed with cardiovascular diseases (CVDs), the leading cause of death globally. The gold standard for identifying these heart problems is via electrocardiogram (ECG). The standard 12-lead ECG is widely used in clinical practice and most of the current research. However, using fewer leads can make ECG more pervasive as it can be integrated with portable or wearable devices. This article introduces two novel techniques to improve the performance of the current deep learning system for 3-lead ECG classification, making it comparable with models that are trained using standard 12-lead ECG. Specifically, we propose a multi-task learning scheme in the form of the number of heartbeats regression and an effective mechanism to integrate patient demographic data into the system. With these two advancements, we got classification performance in terms of F1 scores of 0.9796 and 0.8140 on two large-scale ECG datasets, i.e., Chapman and CPSC-2018, respectively, which surpassed current state-of-the-art ECG classification methods, even those trained on 12-lead data. Our source code is available at github.com/lhkiem28/LightX3ECG.

I. INTRODUCTION

Cardiovascular diseases (CVDs) are a group of disorders of the heart and blood vessels, including coronary heart disease, cerebrovascular disease, rheumatic heart disease, and many other conditions. CVDs are one of the primary causes of death globally, accounting for approximately 19.05 million deaths worldwide in 2020 [1]. Fortunately, premature deaths can be avoided by early detection of people who are most at risk for CVDs and by ensuring they receive the appropriate care with counseling and medications. The electrocardiogram (ECG), which is a representation of the electrical activity of the heart obtained by placing electrodes on the body surface, shows the pathological states of the cardiovascular system in its waveforms or rhythms. Therefore, this electrical signal is a widely used, simple but effective tool to identify cardiovascular abnormalities in patients. The usual structure of an ECG beat is shown in Fig 1, consisting of a P wave, a QRS complex with an R peak, a T wave, and other parts such as PR, QT intervals, or PR, ST segments.

It is estimated that 1.5 million and 3 million electrocardiograms are conducted every single day across the globe [3, 4], this

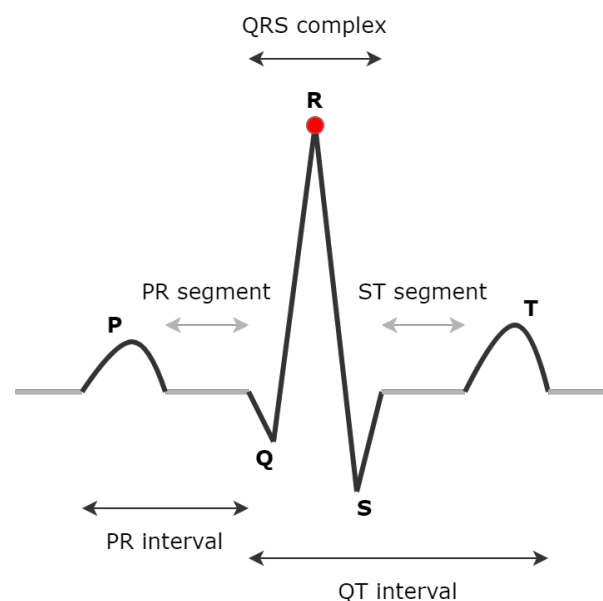


Fig. 1. The usual structure of an ECG beat [2].

makes computer-aided, automatic ECG analysis essential, especially in low- and middle-income countries, where high-quality and experienced cardiologists are extremely scarce. In the past few years, research works concerning automatic ECG analysis have witnessed rapid development, and deep learning-based methods have become the preferred approach. For standard 12-lead ECG, Acharya et al. [5] early developed a 9-layer 1D-CNN to identify 5 different types of cardiovascular abnormalities, Zhang et al. [6] proposed using 1D-ResNet34, Zhu et al. [7] ensemble two 1D-SEResNet34s and one set of expert rules to identify more types of abnormalities. Li et al. [8] combined a sparse Autoencoder and Hidden Markov Model for diagnosing obstructive sleep apnea. Moreover, Rahman et al. [9] tried to early diagnose COVID-19 using ECG trace images. Recently, with the development of many small, low-cost, and easy-to-use ECG-enable devices [10, 11, 12] which provide a subset of ECG leads rather than the entire set, newer methods are being developed to do ECG

* Corresponding author: khiem.lh@vinuni.edu.vn

classification based on reduced-lead ECG rather than standard 12-lead data. For instance, Drew et al. [13] demonstrated that interpolated 12-lead ECG, which is derived from a reduced-lead set (limb leads plus V1 and V5), is comparable to standard 12-lead ECG for diagnosing wide-QRS-complex tachycardias and acute myocardial ischemia. Green et al. [14] also found that the leads III, aVL, and V2 together yielded a similar performance as the full 12-lead ECG for diagnosing acute coronary syndrome. Cho et al. [15] claimed that myocardial infarction could be detected not only with a conventional 12-lead ECG but also with a limb 6-lead ECG. Unlike previous studies, our work uses a combination of only three ECG leads (I, II, and V1) as the input for the system to strike a balance between high classification performance and the ease of signal acquisition, thus providing further support to demonstrate the ability of reduced-lead ECG to identify a wider range of cardiovascular abnormalities [2].

From the perspective of deep learning, there are two primary strategies to enhance the performance of a learning-based model, that are internal improvement and external information integration. For internal improvements, existing research works mainly focus on model architecture changes or robust loss function usages. For instance, Cheng et al. [16] used a much larger kernel size in order to capture longer patterns in ECG signals. Yao et al. [17] constructed a Time-Incremental ResNet18 (TI-ResNet18), a combination of a 1D-ResNet18 and an LSTM network to capture both spatial and temporal features. Zhu et al. [7] suggested using a Sign loss function and Romdhane et al. [18] suggested using a Focal loss function to reduce the negative effects caused by class imbalance problems in datasets. In this work, we propose a novel, simple but effective yet internal improvement as a multi-task learning scheme where we perform a subtask of the number of heartbeats regression. In general, a multi-task learning scheme offers the advantage of reducing overfitting through shared representations, thus improving the system's generalization. Moreover, by performing the proposed heartbeat counting subtask, the system is relatively forced to capture periodic features in ECG signals, which are mostly ignored in other methods.

The idea of utilizing external data or metadata to improve classification performance has been explored in other domains. In particular, Li et al. [19] attempted to enhance landmark identification in tourist photographs by using GPS information. Ellen et al. [20] incorporated both geo-temporal and hydrographic metadata to boost the performance of plankton image classification. In the medical domain, Ningrum et al. [21] observed the use of patient metadata such as age, gender, and anatomical site leads to higher accuracy in malignant melanoma detection on dermoscopic images. In this work, we design an effective encoding process and a shallow neural network to embed patient demographic data, including age and gender, which along with ECG recordings, to further increase the power of the developed ECG classification system.

To summarize, our main contributions are as follows:

- We propose a novel multi-task learning scheme in the form of the number of heartbeats regression to help the system capture periodic features and increase its generalization.
- An effective mechanism is designed to integrate patient demographic data, including age and gender, into the system, thus boosting the classification performance.
- Extensive experiments were conducted on two large-scale multi-lead ECG datasets, i.e., Chapman and CPSC-2018, where the enhanced system significantly outperformed state-of-the-art models that are trained using 12-lead ECG.

II. BASELINE SYSTEM

In this section, we briefly introduce our current system, X3ECG [2]. Firstly, the overall architecture is described. Next, we present the most important part, the Lead-wise Attention module, in more detail.

A. Overall Architecture

We employ three distinct 1D-SEResNets18 [22] as backbones to extract features from three input ECG leads separately. This multi-input strategy is reasonable since these kinds of signals usually require separate treatment. Then, a novel Lead-wise Attention module is designed as the aggregation technique to explore the most essential input lead and merge outputs of these backbones, resulting in a more robust representation that is then sent through a Fully-Connected (FC) layer to perform classification.

B. Lead-wise Attention

To achieve an end-to-end classification system, the outputs, also known as features, extracted from backbones, must be combined. Typically, one can combine these features by simply applying a summation or concatenation operation to them, but this is usually ineffective due to their simplicity. Thus, we propose a Lead-wise Attention module to more effectively ensemble these features together and acquire a final robust feature which is then routed to the last FC layer, the classifier, to perform classification. Our Lead-wise Attention module is described in Fig 2.

Firstly, features from our backbones are concatenated and sent through a sequential list of layers including an FC, a BatchNorm, a Dropout, followed by another FC layer and a Sigmoid function to determine the attention score, or importance score for each feature. Subsequently, the final feature is obtained by taking a weighted sum over these features by corresponding generated scores. This module can be formulated as:

$$\mathbf{f}_{\text{merged}} = \sum_{i=1}^3 \alpha_i \mathbf{f}_i, \quad (1)$$

$$\alpha = \text{Sigmoid}(\text{FC}(\text{FC}(\text{Concat}[\mathbf{f}_i | i = \overline{1, 3}]))).$$

The BatchNorm, Dropout layer, and ReLU function are ignored in the above formulation for simplicity.

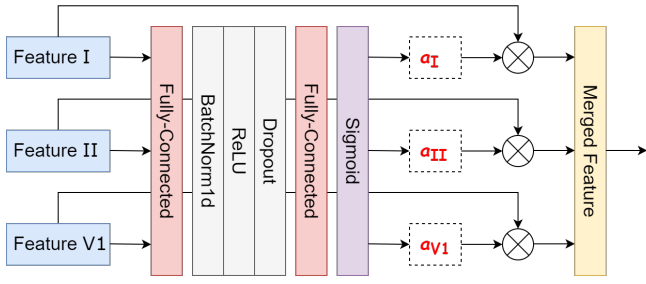


Fig. 2. The proposed Lead-wise Attention module.

III. PROPOSED TECHNIQUES

In this section, we sequentially describe the proposed internal improvement, namely heartbeat counting (HC) and the mechanism of demographic data integration (DDI), that form the final enhanced system, as illustrated in Fig 3.

A. Heartbeat Counting

We formulate the heartbeat counting subtask as a simple regression problem. Specifically, an FC layer with one output unit is used as an auxiliary regressor, taking the merged feature from the Lead-wise Attention module f_{merged} as input and trying to predict how many heartbeats there are in ECG recording:

$$n_{\text{pred}} = \text{FC}(f_{\text{merged}}) \quad (2)$$

At the training stage, the number of heartbeats of each ECG recording is required as ground truth. But the majority of datasets lack this information. Alternatively, we utilize a popular and highly accurate tool called NeuroKit2 [23] to detect R peaks in lead I ECG signal of a recording. The number of R peaks represents the number of heartbeats, thus is considered ground truth. A Mean Absolute Error (MAE) is employed to compute the difference between ground truth n_{gt} and the predicted number of heartbeats n_{pred} . Finally, the whole system is trained end-to-end with a combined loss function as:

$$\mathcal{L}_{\text{combined}} = \mathcal{L}_{\text{cls}} + \lambda |n_{\text{gt}} - n_{\text{pred}}|, \quad (3)$$

where \mathcal{L}_{cls} is the main CVDs classification loss function, and λ is a hyperparameter, which controls the contribution and effect of the auxiliary regressor on the whole system. At the inference stage, the auxiliary branch will be ignored, thus there is no additional cost to the system.

B. Demographic Data Integration

Different research has demonstrated that demographic data, including age and gender, is associated with cardiovascular risks. Also, cardiologists need to additionally use this information to make final decisions [24, 25]. Based on that, we leverage this external data, available in most ECG datasets to enhance the system's performance further.

1) Feature Encoding:

We have age and gender features in the form of discrete and categorical features, respectively, that need an effective encoding strategy before feeding into the system. Firstly, the age feature is converted into a categorical feature by being divided into 7 groups including 0 to 4 years, 5 to 14 years, 15 to 22 years, 23 to 41 years, 42 to 56 years, 57 to 68 years, and over 69 years that are respectively corresponding to early childhood, childhood, adolescence, youth, maturity, old age, and late old age. Then, we apply one-hot encoding to these two categorical features and treat missing values as their own category. Finally, we obtain a one-hot feature vector with a length of 11, which can be used by the system.

2) Feature Fusion:

To achieve an end-to-end system, we design a shallow neural network, also known as Multi-layer Perceptron (MLP) to process the encoded demographic data and fuse the output feature with the merged feature from the Lead-wise Attention module f_{merged} by a simple concatenation operation.

Algorithm 1 PyTorch-like pseudocode for the Feature Fusion

```

MLP = nn.Sequential(
    nn.Linear(11, 128),
    nn.BatchNorm1d(128),
    nn.ReLU(),
    nn.Dropout(0.2),
    nn.Linear(128, 128),
    nn.BatchNorm1d(128),
    nn.ReLU(),
)
f_demog = MLP(demog)
f_final = torch.cat([f_demog, f_merged])

```

Algorithm 1 provides the pseudo-code for our designed MLP architecture and the fusion operation in a PyTorch-like style [26]. The final feature is used for performing the main task CVDs classification.

IV. EXPERIMENTS AND RESULTS

In this section, we evaluate the proposed techniques on two large-scale ECG datasets, i.e., Chapman and CPSC-2018, with class frequency and demographic data statistics shown in Table I. Extensive experiments were performed to compare our system with other ECG classification methods and to validate the contributions of each component in the whole system.

A. Datasets

Chapman [27]. Chapman University and Shaoxing People's Hospital collaborated to establish this large-scale *multi-class* ECG dataset which consisted of 10,646 12-lead ECG recordings. Each recording was taken over 10 seconds with a sampling rate of 500 Hz and labeled with 11 common diagnostic classes. The amplitude unit was microvolt. These 11 classes were hierarchically grouped into 4 categories including AFIB, GSVT, SB, and SR.

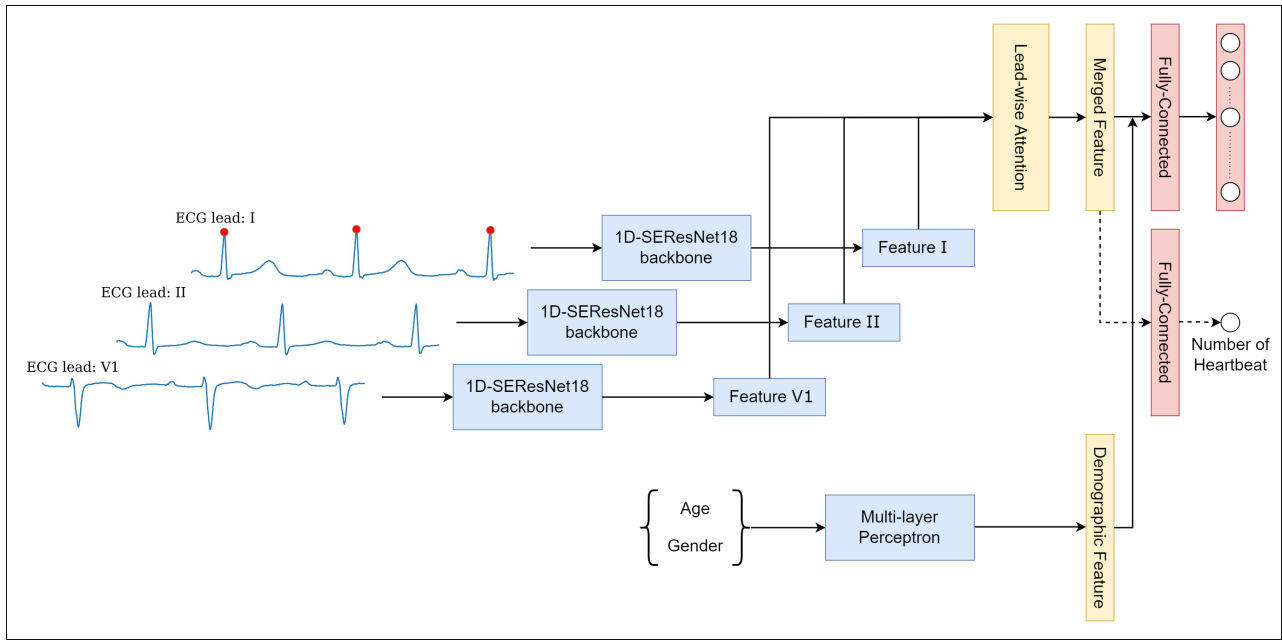


Fig. 3. An overview of the whole proposed system.

Table I
DESCRIPTION OF TWO DATASETS.

Chapman			
Class	Frequency (%)	Male (%)	Age
AFIB	2225 (20.90)	1298 (58.34)	72.90 ± 11.68
GSVT	2307 (21.67)	1152 (49.93)	55.44 ± 20.49
SB	3889 (36.53)	2481 (63.80)	58.34 ± 13.95
SR	2225 (20.90)	1025 (46.07)	50.84 ± 19.25
CPSC-2018			
Class	Frequency (%)	Male (%)	Age
NSR	918 (13.35)	363 (39.54)	41.56 ± 18.45
AF	1221 (17.75)	692 (56.67)	71.47 ± 12.53
IABV	722 (10.50)	490 (67.87)	66.97 ± 15.67
LBBB	236 (03.43)	117 (49.58)	70.48 ± 12.55
RBBB	1857 (27.00)	1203 (64.78)	62.84 ± 17.07
PAC	616 (08.96)	328 (53.25)	66.56 ± 17.71
PVC	700 (10.18)	357 (51.00)	58.37 ± 17.90
STD	869 (12.64)	252 (29.00)	54.61 ± 17.49
STE	220 (03.20)	180 (81.82)	52.32 ± 19.77

CPSC-2018 [28]. The first China Physiological Signal Challenge was held in 2018, and a large-scale *multi-label* ECG dataset was made accessible to the public. This dataset contained 6.877 12-lead ECG recordings with a sampling rate of 500 Hz, an amplitude unit of millivolt, and durations ranging from 6 to 60 seconds. These ECG recordings were labeled with 9 diagnostic classes including NSR, AF, IABV, LBBB, RBBB, PAC, PVC, STD, and STE.

B. Implementation Details

As a deep learning system requires inputs to be of the same length, all ECG recordings were fixed at 10 seconds in length in both datasets. This was done by truncating the part exceeding the first 10 seconds for longer recordings and padding shorter ones with zero. Then, signals were filtered using a zero-phase method with 3rd order Butterworth bandpass filter with a frequency band from 1 Hz to 47 Hz. We took leads I, II, and V1 from each ECG recording to construct the input with the shape of 3x5000 and fed it into the system.

For evaluation, we applied a 10-fold cross-validation strategy following some previous works [6, 27]. We stratify divided each of the two datasets into 10 folds and performed 10 rounds of training and evaluation. At each round, 8 folds; 1 fold; and 1 remaining fold were used as training, validation, and test set, respectively. In the multi-label classification manner, the optimal threshold of each class was searched in a range (0.05, 0.95) with a step of 0.05 to achieve the best F1 score on the validation set. We report the average performance of 10 rounds on the test set in terms of F1 score and accuracy. In all experiments on our system, the system was optimized by an Adam optimizer with an initial learning rate of 1e-3 and a weight decay of 5e-5 for 70 epochs. We used the Cosine Annealing scheduler in the first 40 epochs to reschedule the learning rate to 1e-4 and then kept it constant in the last 30 epochs. Cross-entropy and binary cross-entropy were utilized as classification loss functions in multi-class and multi-label manners, respectively. When applying the proposed heartbeat counting technique, the control hyperparameter λ was set to 0.02. All experiments were performed on a workstation with an NVIDIA GeForce RTX 3090 TURBO 24G.

Table II
COMPARISON OF THE WHOLE PROPOSED SYSTEM TO SOTA METHODS.

Method	Chapman macro-F1	Chapman Acc	CPSC-2018 macro-F1	CPSC-2018 Acc
TI-ResNet18 [17]	0.9348	0.9399	0.7387	0.9484
InceptionTime [29]	0.9640	0.9671	0.7908	0.9604
1D-SEResNet34 [6]	0.9695	0.9723	0.8036	0.9635
1D-SEResNet34 w/ Focal loss [18]	0.9722	0.9738	0.8095	0.9640
X3ECG [2]	0.9746	0.9775	0.8025	0.9637
X3ECG w/ HC	0.9783	0.9803	0.8110	0.9648
X3ECG w/ HC + DDI	0.9796	0.9817	0.8140	0.9652

C. Comparison with SOTA Methods

We compared the whole proposed system with state-of-the-art ECG classification methods including TI-ResNet18 [17], InceptionTime [29], 1D-SEResNet34 [6], and 1D-SEResNet34 with Focal loss [18]. While our system was trained on only three ECG leads, all other methods were trained using standard 12-lead ECG. From Table II, X3ECG enhanced with the HC technique helped us to set new state-of-the-art performances on two datasets, outperforming other methods in terms of F1 score and accuracy. Furthermore, the system enhanced with both two techniques, HC and DDI, boosted these performances further. In particular, we got F1 scores of 0.9796 and 0.8140, and accuracies of 0.9817 and 0.9652 on two datasets, Chapman and CPSC-2018 respectively.

V. CONCLUSION

In this article, we introduced an accurate deep learning system that uses a combination of three ECG leads (I, II, and V1) to identify cardiovascular abnormalities. Besides that, two novel techniques heartbeat counting and demographic data integration, which were developed to enhance system performance were presented. The whole proposed system surpassed current state-of-the-art ECG classification methods even those trained on standard 12-lead ECG. Also, we emphasize that heartbeat counting and demographic data integration are plug-in techniques and can be easily adapted to other different models that can accelerate other research.

VI. ACKNOWLEDGMENT

The authors would like to thank the VinUni-Illinois Smart Health Center (VISHC) and the source of funding: VinUni Seed Grant 2020. For more information, please visit the website at vinuni.edu.vn/vinuni-research/seed-funding-program/. Besides, this work was also funded by Vingroup Joint Stock Company (Vingroup JSC), and supported by Vingroup Innovation Foundation (VINIF) under project code VINIF.2021.DA00128.

REFERENCES

- [1] Connie W. Tsao, Aaron W. Aday, Zaid I. Almarzooq, Alvaro Alonso, Andrea Z. Beaton, and Marcio S. Bittencourt et al. Heart disease and stroke statistics. *Circulation*, 145(8):e153–e639, 2022.
- [2] Khiem H. Le, Hieu H. Pham, Thao BT. Nguyen, Tu A. Nguyen, Tien N. Thanh, and Cuong D. Do. Lightx3ecg: A lightweight and explainable deep learning system for 3-lead electrocardiogram classification. *pre-print arXiv:2207.12381*, 2022.
- [3] Paul Kligfield, Leonard S Gettes, James J Bailey, Rory Childers, Barbara J Deal, E William Hancock, Gerard van Herpen, Jan A Kors, Peter Macfarlane, David M Mirvis, Olle Pahlm, Pentti Rautaharju, Galen S Wagner, American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology, and et al. Recommendations for the standardization and interpretation of the electrocardiogram: part i: The electrocardiogram and its technology. *Circulation*, 115(10):1306–1324, February 2007.
- [4] Annemarijn Sm Steijlen, Kaspar Mb Jansen, Armagan Albayrak, Derk O Verschure, and Diederik F Van Wijk. A novel 12-lead electrocardiographic system for home use: Development and usability testing. *JMIR Mhealth Uhealth*, 6(7):e10126, July 2018.
- [5] U Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, Muhammad Adam, Arkadiusz Gertych, and Ru San Tan. A deep convolutional neural network model to classify heartbeats. *Computers in Biology and Medicine*, 89, 08 2017.
- [6] Dongdong Zhang, Samuel Yang, Xiaohui Yuan, and Ping Zhang. Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram. *iScience*, 24(4):102373, 2021.
- [7] Zhaowei Zhu, Xiang Lan, Tingting Zhao, Yangming Guo, Pipin Kojodjojo, Zhuoyang Xu, Zhuo Liu, Siqi Liu, Han Wang, Xingzhi Sun, and Mengling Feng. Identification of 27 abnormalities from multi-lead ecg signals: an ensemble se-resnet framework with sign loss function. *Physiological Measurement*, 42, 06 2021.
- [8] Kunyang Li, Pan Weifeng, Qing Jiang, and Guan-Zheng Liu. A method to detect sleep apnea based on deep neural network and hidden markov model using single-lead ecg signal. *Neurocomputing*, 294, 03 2018.
- [9] Tawsifur Rahman, Alex Akinbi, Muhammad Chowdhury,

- Tarik Rashid, Abdulkadir Sengur, Amith Khandakar, Khandaker Islam, and aras masood. Cov-ecgnet: Covid-19 detection using ecg trace images with deep convolutional neural network. *Health Information Science and Systems*, 10, 01 2022.
- [10] Mintu P. Turakhia, Donald D. Hoang, Peter J Zimetbaum, Jared D. Miller, Victor F. Froelicher, Uday N. Kumar, Xiangyan Xu, Felix Yang, and Paul A. Heidenreich. Diagnostic utility of a novel leadless arrhythmia monitoring device. *The American journal of cardiology*, 112 4:520–4, 2013.
- [11] John Chorba, Avi Shapiro, Le An, John Maidens, John Prince, Steve Pham, Mia Kanzawa, Daniel Barbosa, Caroline Currie, Catherine Brooks, Brent White, Anna Huskin, Jason Paek, Jack Geocarlis, Dinatu Elnathan, Ria Ronquillo, Roy Kim, Zenith Alam, Vaikom Mahadevan, and James Thomas. Deep learning algorithm for automated cardiac murmur detection via a digital stethoscope platform. *Journal of the American Heart Association*, 10:e019905, 04 2021.
- [12] Gregory Marcus. The apple watch can detect atrial fibrillation: so what now? *Nature Reviews Cardiology*, 17:1–2, 12 2019.
- [13] Barbara Drew, Michele Pelter, Donald Brodnick, Viapn Yadav, Debbie Dempel, and Mary Carey. Comparison of a new reduced lead set ecg with the standard ecg for diagnosing cardiac arrhythmias and myocardial ischemia. *Journal of electrocardiology*, 35 Suppl:13–21, 10 2002.
- [14] Michael Green, Mattias Ohlsson, Jakob Forberg, Jonas Björk, Lars Edenbrandt, and Ulf Ekelund. Best leads in the standard electrocardiogram for the emergency detection of acute coronary syndrome. *Journal of electrocardiology*, 40:251–6, 08 2007.
- [15] Younghoon Cho, Joon-Myoung Kwon, Kyung-Hee Kim, Jose Medina-Inojosa, KI-Hyun Jeon, Soohyun Cho, Soo Lee, Jinsik Park, and Byung-Hee Oh. Artificial intelligence algorithm for detecting myocardial infarction using six-lead electrocardiography. *Scientific Reports*, 10, 11 2020.
- [16] Jinyong Cheng, Qingxu Zou, and Yunxiang Zhao. Ecg signal classification based on deep cnn and bilstm. *BMC Medical Informatics and Decision Making*, 21, 12 2021.
- [17] Qihang Yao, Xiaomao Fan, Yunpeng Cai, Ruxin Wang, Liyan Yin, and Ye Li. Time-incremental convolutional neural network for arrhythmia detection in varied-length electrocardiogram. In *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing*, pages 754–761, 2018.
- [18] Taissir Fekih Romdhane, Haikel Alhichri, Ridha Ouni, and Mohamed Atri. Electrocardiogram heartbeat classification based on a deep convolutional neural network and focal loss. *Computers in Biology and Medicine*, 123:103866, 2020.
- [19] Yunpeng Li, David J. Crandall, and Daniel P. Huttenlocher. Landmark classification in large-scale image collections. In *2009 IEEE 12th International Conference on Computer Vision*, pages 1957–1964, 2009.
- [20] Jeffrey Ellen, Casey Graff, and Mark Ohman. Improving plankton image classification using context metadata. *Limnology and Oceanography: Methods*, 17, 07 2019.
- [21] Dina Nur Anggraini Ningrum, Sheng-Po Yuan, Woon-Man Kung, Chieh-Chen Wu, I-Shiang Tzeng, Chu-Ya Huang, Jack Yu-Chuan Li, and Yao-Chin Wang. Deep learning classifier with patient’s metadata of dermoscopic images in malignant melanoma detection. *J Multidiscip Healthc*, 14:877–885, April 2021.
- [22] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *2018 Conference on Computer Vision and Pattern Recognition*, pages 7132–7141, 2018.
- [23] Dominique Makowski, Tam Pham, Zen J. Lau, Jan C. Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and S. H. Annabel Chen. NeuroKit2: A python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 53(4):1689–1696, 2021.
- [24] Jennifer L Rodgers, Jarrod Jones, Samuel I Bolleddu, Sahit Vanthenapalli, Lydia E Rodgers, Kinjal Shah, Krishna Karia, and Siva K Panguluri. Cardiovascular risks associated with gender and aging. *J Cardiovasc Dev Dis*, 6(2), April 2019.
- [25] P Jousilahti, E Vartiainen, J Tuomilehto, and P Puska. Sex, age, cardiovascular risk factors, and coronary heart disease: a prospective follow-up study of 14 786 middle-aged men and women in finland. *Circulation*, 99(9):1165–1172, March 1999.
- [26] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.
- [27] Jianwei Zheng, Jianming Zhang, Sidy Danioko, Hai Yao, Hangyuan Guo, and Cyril Rakovski. A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. *Scientific Data*, 7, 02 2020.
- [28] Feifei Liu, Chengyu Liu, Lina Zhao, Xiangyu Zhang, Xiaoling Wu, Xiaoyan Xu, Yulin Liu, Caiyun Ma, Shoushui Wei, Zhiqiang He, Jianqing Li, and Eddie Ng. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. *Journal of Medical Imaging and Health Informatics*, 8:1368–1373, 09 2018.
- [29] Hassan Ismail Fawaz, Benjamin Lucas, Germain Forestier, Charlotte Pelletier, Daniel F. Schmidt, Jonathan Weber, Geoffrey I. Webb, Lhassane Idoumghar, Pierre-Alain Muller, and François Petitjean. Inceptiontime: Finding alexnet for time series classification. *Data Min. Knowl. Discov.*, 34(6):1936–1962, nov 2020.