Computers in Biology and Medicine 127 (2020) 104077

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

Deep learning for digitizing highly noisy paper-based ECG records

Yao Li [a,b,c,1], Qixun Qu [d,1], Meng Wang [d], Liheng Yu [e], Jun Wang [d,f,g], Linghao Shen [d,**], Kunlun He [b,c,*]

[a] Medicine School of Chinese PLA, Beijing, 100853, China
[b] Core Laboratory of Translational Medicine, Chinese PLA General Hospital, Beijing, 100853, China
[c] Beijing Key Laboratory of Chronic Heart Failure Precision Medicine, Chinese PLA General Hospital, Beijing, 100853, China
[d] Shenzhen Digital Life Institute, Shenzhen, Guangdong, 518000, China
[e] The 980th Hospital of PLA Joint Logistical Support Force (Bethune International Peace Hospital), Shijiazhuang, Hebei, 050082, China
[f] ICarbonX, Zhuhai, Guangdong, 519000, China
[g] Macau University of Science and Technology, Macau

ARTICLE INFO

Keywords:
Electrocardiogram
Digitization
Deep learning
Image segmentation
Signal processing

ABSTRACT

Electrocardiography (ECG) is essential in many heart diseases. However, some ECGs are recorded by paper, which can be highly noisy. Digitizing the paper-based ECG records into a high-quality signal is critical for further analysis. We formulated the digitization problem as a segmentation problem and proposed a deep learning method to digitize highly noisy ECG scans. Our method extracts the ECG signal in an end-to-end manner and can handle different paper record layouts. In the experiment, our model clearly extracted the ECG waveform with a Dice coefficient of 0.85 and accurately measured the common ECG parameters with more than 0.90 Pearson's correlation. We showed that the end-to-end approach with deep learning can be powerful in ECG digitization. To the best of our knowledge, we provide the first approach to digitize the least informative noisy binary ECG scans and potentially be generalized to digitize various ECG records.

1. Introduction

Electrocardiogram (ECG) is non-invasive, continental, and inexpensive and is currently one of the most commonly used diagnostic tools for cardiovascular diseases. Some typically altered ECG signals point to exact diseases, including coronary artery diseases and arrhythmia, which provide useful information for diagnosis and prognosis [1–3]. With the development of digitization, most ECGs are recorded and stored in both digital and paper formats [4]. The digitized ECG can be quickly, simultaneously measured and automatically interpreted, and complex ECG parameters can be measured [5]. When the ECG is old and recorded only on paper reports, the scanning and digitization process is required. However, the defects of the scanned digitized ECG prevent direct waveform extraction from scans. Excessive noise and interference, patient information irrelevant to the ECG signal itself, and manual marking on the paper will affect the segmentation and analysis of the ECG. Some studies have developed denoising methods [6–8] for different situations; however, no single method is universal [9,10].

Existing ECG digitization methods generally have 4 steps (Fig. 1A): ROI (**R**egion-**O**f-**I**nterest) detection, grid removal, signal refinement, and vectorization [10]. The ROI detection extracts the rough ECG waveform region to bypass the noise in an unrelated region; then, the grid removal further removes the background grid from the ECG scan and retains the clean waveform. Without manual operation such as [11], these two steps highly depend on the quality and layout of the ECG scan. For example [12], considered the RGB scan (Fig. 2A), where the signal and background can be robustly distinguished by different colours [13]. considered a more practical grey-level scan (Fig. 2B and C), where Otsu's thresholding [14] was used to separate the signal and background. Nonetheless, there are many ECG records in medical reports, which are binary and highly noisy in both signal waveform and grid. Signal refinement is another challenge, since printing and scanning may distort both signal and background [12]. removed the salt-and-pepper noise by median filtering and [13] filtered out small connected objects by connectivity detection [15]. Continuity algorithms were also generally applied [16] to interpolate waveform gaps. However, errors may

* Corresponding author. Core Laboratory of Translational Medicine, Chinese PLA General Hospital, Beijing, 100853, China.
** Corresponding author.
E-mail addresses: shenlinghao@icarbonx.com (L. Shen), kunlunhe@plagh.org (K. He).
[1] Yao Li and Qixun Qu joint first authors and contributed equally to the article.

https://doi.org/10.1016/j.compbiomed.2020.104077
Received 15 June 2020; Received in revised form 19 October 2020; Accepted 20 October 2020
Available online 28 October 2020

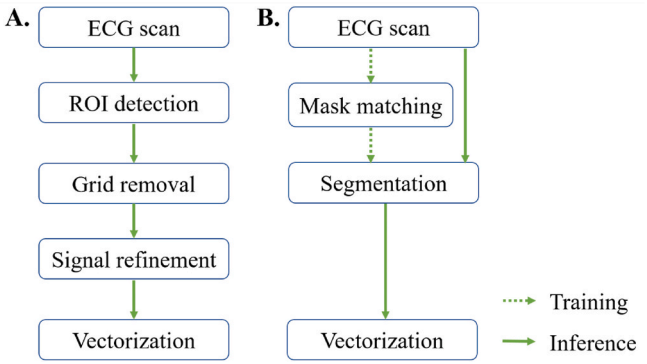

Fig. 1. ECG digitization procedures. **A.** Ordinary image processing procedures for digitization. Each step relies on the result of the previous step, and errors may accumulate. **B.** Deep learning approach for digitization; break lines represent training, and solid lines represent inference. The segmentation is an end-to-end pixelwise labelling that resembles the intermediate steps in A.

accumulate in each step, especially for the complex but practical cases, as shown in Fig. 2D. In summary, existing digitization techniques employ human-designed rules and hand-crafted features, which can make it difficult to tackle more complex and diversified situations.

To overcome these difficulties, we propose a general deep learning framework to handle ROI detection, grid removal, and signal refinement in an end-to-end manner (the code and trained model are available in the supplementary *code_and_model.rar*). The digitization problem can be treated as a pixelwise labelling problem, where either the signal or background should be assigned to the pixels. This is a typical image segmentation problem [15], and we propose to solve this problem by using the U-Net [17] deep neural network. The end-to-end approach removes the dependence on hand-craft features and enables overall optimization. Manually annotating the waveforms on ECG scans can be troublesome; thus, template matching [18] was used to robustly retrieve the segmentation mask from their corresponding digital signals. The training data were also augmented by adding both structural and random noise to the true ECG signal. The U-Net model processes the image in a patch-by-patch manner, which is unaware of the signal layout. Thus, different layouts of ECG scans can be naturally automatically handled. After segmentation, the clean ECG waveform can be extracted and vectorized based on the scale of the grid. In the experiment, we attempted to address the following questions for the proposed framework:

1. Is U-Net better than other common deep learning segmentation models for ECG waveform extraction?
2. Is the proposed image segmentation framework better than existing methods for ECG digitization?
3. Can the trained segmentation model extract the ECG waveforms in another type of ECG scan?

The waveform extraction was measured by the segmentation performance, where the U-Net provided superior results. We also validated the proposed method in terms of signal similarity and common ECG parameters, which resulted in good performance.

In this manuscript, we formulated ECG digitization as an image segmentation problem and provided a deep learning solution. The proposed method worked for highly noisy and binary ECG scans, which failed many existing available methods. Experimentally, we showed that our approach could be generalized to digitize different types of ECG records. We believe that the proposed framework can benefit the development of ECG algorithms and their application to real-world ECG data.

2. Material and methods

2.1. Data collection

All ECGs were from the Department of Medical Record Management, the First Medical Centre, Chinese People's Liberation Army General

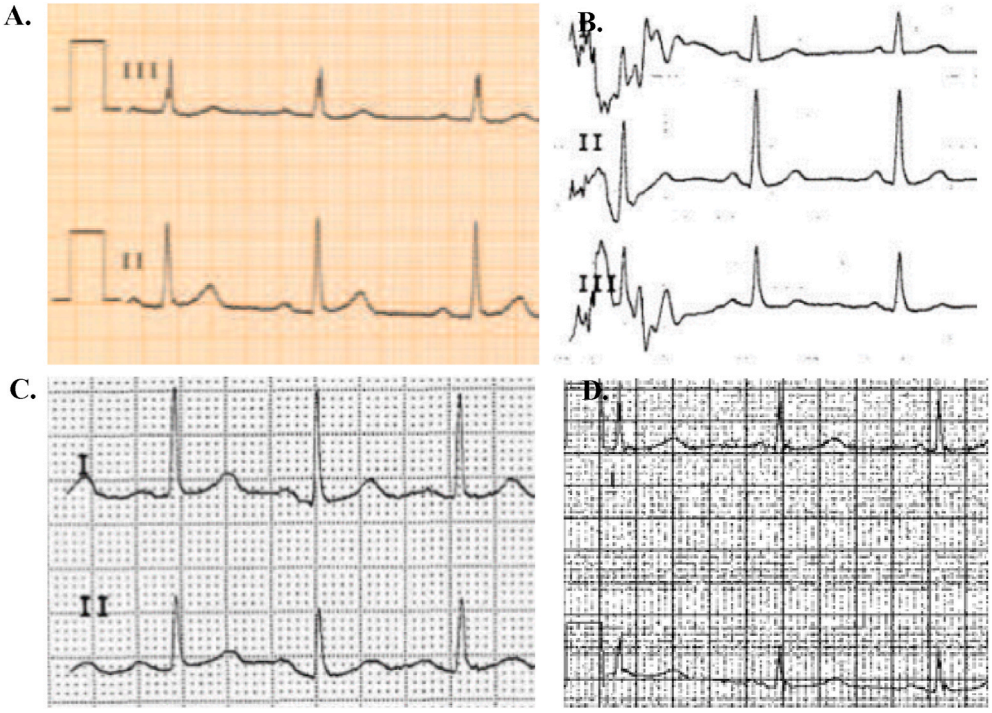

Fig. 2. ECG scans at different levels of noise. **A.** Colourful scan, grid and signal can be directly distinguished by colour (taken from Ref. [12]). **B**, **C.** Grey-level scan with (C) or without (B) uniform grid (taken from Ref. [13]); the ECG waveform can be extracted by thresholding or histogram methods. **D.** Binary scan with uneven grid (taken by scanning the ECG on a medical report), which cannot be processed by previous methods but is handled by the proposed method.

Hospital. The ECGs were recorded in paper and digital format by specifically trained and experienced technicians with GE 5500 (General Electric Company, US). The paper speed was 50 mm/s, and the calibration was 10 mm/mV. Three leads (limb leads, augmented limb leads, V1–V3, V4–V6) were simultaneously recorded, and every lead recording was approximately 2.5 s long. The digital ECGs were immediately stored in electronic health records. All ECGs on paper reports were scanned using HP LaserJet Enterprise M606 N (Hewlett-Packard, US) in TIFF format and stored in electronic health records.

In total, a dataset of 96 ECG signals with both paper scans and digital ground truth was used in this study (illustrated in Fig. 4A and B), which contained binary ECG scans of resolution 886 × 1120 pixels. Meanwhile, the dataset contains three different ECG scan layouts (Fig. S1): 81 four-line, 14 six-line, and one seven-line ECG scans. For the four-line and six-line scans, the first three lines are the 12 ECG leads; for the seven-line scan, the first six lines are the 12 ECG leads. The last few lines in all scans are long-term measurements, where all three layouts contain a long V1 lead. All scans were manually de-identified to prevent privacy leakage.

2.2. Training data preparation

To train the image segmentation model, ECG scans with waveform annotation are necessary. However, matching the ground truth digital signal to the waveforms on the ECG scan is not trivial, since the position of each lead on different scans can vary. Additionally, the waveform on paper may not perfectly match the corresponding digital signal, since the paper may slightly stretch or shrink. These problems may be solved by considering the best matching among variations. The signal is transformed to a 2-D template by one-hot encoding the signal value in each time step. Formally, given the ECG scan image I and signal template S, we find the best transform $\boldsymbol{H}$ and location (x,y) that minimize:

$$R(\boldsymbol{H},x,y)=\sum_{x',y'}[\boldsymbol{H}(S)(x',\ y')\ -I(x+x',y+y')]^2,$$

where $\boldsymbol{H}$ includes amplifying/suppressing and up/down-sampling at different scales. This method was applied to each lead, and the segmentation mask of the waveform was drawn based on the best $\boldsymbol{H}(S)$ at the best (x, y). This template-matching approach produces clean and continuous waveforms. Therefore, it enables the segmentation model to learn the denoising and waveform interpolation in an end-to-end manner. The raw images were scaled from 0–255 to 0–1 and colour-inverted for conventional representation (0 for background; 1 for waveform).

To utilize the digital ECG signals without the corresponding scans, we generalized some pseudo-noisy ECG scans. First, for each signal, all leads were plotted on a binary image according to the common scale and layout. Then, both structural noise (thick and thin grids; thickness of the waveform) and random noise (pepper-and-salt noise) at different strengths were added (Fig. S2). In total, 4 signals were used, and 24 pseudo-scans were generated. Since the augmented training data were still far from perfect, we only used them to pre-train the U-Net model.

2.3. Proposed framework for ECG digitization

The ECG digitization depends on the waveform extraction from scans, where each pixel should be classified as background or waveform (denoted as $y \in \{0,1\}$, respectively). Notably, the intermediate steps (ROI detection, grid removal and signal refinement) in Fig. 1A can also be treated as step-by-step pixel labelling. More specifically, the ROI detection labels the pixel outside the focused region as background to narrow down the candidate pixels. Then, the grid removal detects the grid pixels and labels them as background. Finally, the signal refinement conducts fine-grained pixel labelling to distinguish between waveforms and backgrounds based on the signal shape, connectivity, noise assumption, etc. From this viewpoint, the proposed method is a learnable and end-to-end reassembly of existing procedures, which ensures its availability and potentiality.

Formally, this problem can be addressed as an image segmentation problem $\boldsymbol{y} = f_\theta(\boldsymbol{x})$. The proposed framework (Fig. 3) uses U-Net as the segmentation model f_θ due to its good records in many medical image segmentation tasks [19] and superior performance in this task (see section 3). To perform segmentation at high resolution, $\boldsymbol{x}$ is a 128 × 128-pixel patch sequentially taken from the entire ECG scan [20,21]. showed that adding residual connections and attention modules could improve the U-Net performance. Thus, the plain convolution in the original U-Net was changed to ResNet [22] blocks with a CBAM [23] module. The ResNet block contains 3 × 3 convolution, batch normalization [24] and LeakyReLU [25] non-linearity for all convolutions. The 2 × 2 stride was used instead of max-pooling for each encoder stage, which empirically could increase the segmentation performance [26]. The model uses the following loss function:

$$L = L_{Dice} + L_{focal},$$

where L_{Dice} is Dice loss [27], and L_{focal} is the focal loss [28] to compensate for the imbalanced numbers of background and waveform pixels.

In the training stage, the patches were randomly sampled (from both the ECG scan and waveform mask) within a 128-pixel range of waveform points to ensure that both background and signal were included. The model parameters were first initialized by 24 pseudonoisy scans (described in section 2.2) by training for a single epoch. Then, the network was trained with batch size 64 for 100 epochs, and Adabound [29] optimizer with an initial learning rate of 1e-3 was used to update the network weights. To utilize the data, a 5-fold cross-validation was conducted for model evaluation. Within each round, the images in the leave-out set (20%) were only used for testing, and the patches were generated from the images in the 80% training set. This procedure ensured that no testing image was leaked to the training set.

In the inference stage, patches were sequentially taken from the ECG scan in the horizontal and vertical directions with a 64-pixel overlap. The final segmentation result combined the output of patches by taking the average at each overlapping pixel, and the threshold of being waveform was 0.15. After the clean ECG waveform was produced, a previous method [13] could be applied to vectorize the ECG signal.

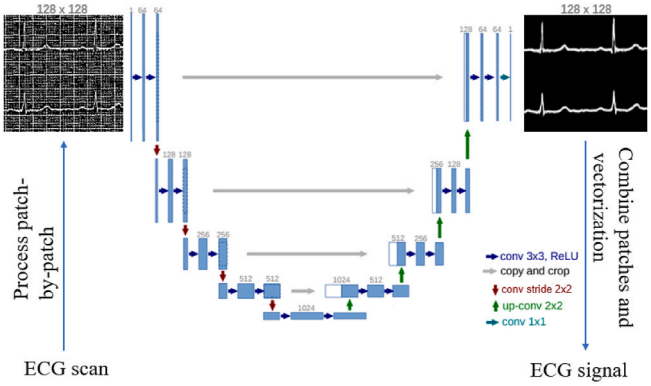

Fig. 3. Proposed framework, which is the architecture of U-Net, from Ref. [17]. Patches are taken from the entire ECG scan and go through the segmentation network to produce a clean waveform. Then, waveforms in each patch are combined (average if overlapping), and vectorization on clean ECG waveform images is performed to obtain the final ECG signal.

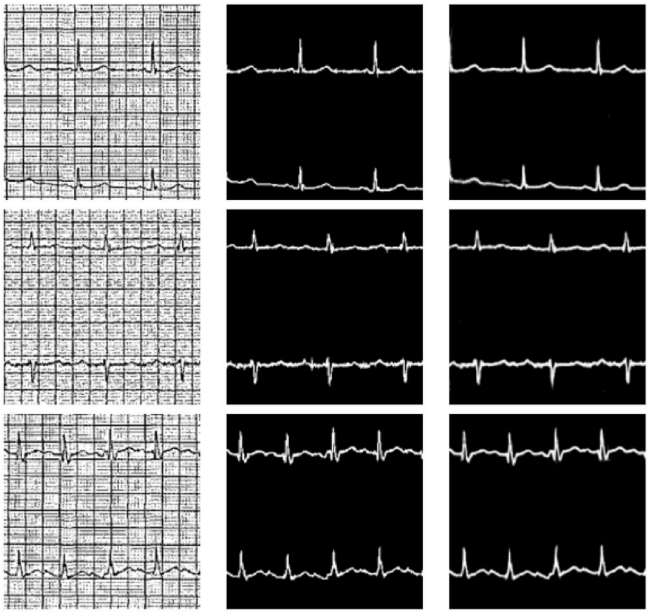

Fig. 4. Example of ECG patches and their corresponding segmentation results in the testing. **Left**: original scans; **middle**: ground truth waveform masks; **right**: segmentation results. The original scans are binary images, where black denotes both signal and grid. The waveform masks and segmentation results are also binary, where bright denotes the signal.

3. Results

3.1. U-net obtained the best waveform extraction performance among the common networks

We compared the waveform extraction performance (examples in Fig. 4C) for widely used segmentation models, including U-Net, FCN [30] and Deeplab-v3 [26]. For a fair comparison, we used ResNet (ResNet18, ResNet34, ResNet50 and ResNet101) as the backbone network for FCN and Deeplab-v3. Since the segmentation was based on patches, different scan layouts were measured together. The segmentation performance was measured based on Dice coefficient [31]. The results are shown in Fig. 5A, and U-Net achieved the best segmentation performance with Dice coefficient = 0.85.

3.2. The proposed method performed well when existing tools failed

The digitized signals were compared to the original signal in terms of signal correlation and ECG parameter measurements. We surveyed the existing tools from Physionet (https://archive.physionet.org/physiotools/digitizing/) and the search engine, but they were unavailable or unable to process binary scans (see supplementary). This fact may indicate the weakness of hand-crafted features, which can be improved by the proposed approach. We used Pearson's correlation coefficient (PCC) to measure the signal and parameter concordance between the digitized signal and the original signal. We measured the correlations of the signal and 5 clinical parameters. The standard deviation and confidence interval were generated by 1000 bootstrap trials. The results are summarized in Table 1.

Overall, the signals demonstrated high concordance (PCC >0.9), and the method accurately estimated both HR and RR (PCC >0.9). The six-line scan also achieved acceptable (PPC >0.8) performance for HRV and QRS. All measured parameters showed significant correlations with the ground truth (lower 95% CI > 0). Notably, the four-line scan had better signal concordance than the six-line scan, but the parameter correlations were lower (the seven-line scan was not considered solely due to insufficient sample). We think that the reason was the short signal length in the four-line scans, where the number of peaks was not

Table 1
Summary of the comparison between digitized signals and original signals. **SD**: standard deviation; **95% CI**: lower and upper bound of 95% confidence interval. The heart rate variability (**HRV**) was computed as the standard deviation of the RR interval.

	PCC	Four-line scan	Six-line scan	Overall
Signal	Mean	0.921	0.900	0.916
	SD	0.008	0.013	0.007
	95% CI	0.902, 0.935	0.871, 0.924	0.901, 0.929
HR	Mean	0.939	0.991	0.949
	SD	0.025	0.006	0.019
	95% CI	0.881, 0.977	0.988, 0.999	0.904, 0.979
RR	Mean	0.962	0.995	0.968
	SD	0.012	0.003	0.009
	95% CI	0.934, 0.982	0.988, 0.999	0.948, 0.984
HRV	Mean	0.667	0.861	0.698
	SD	0.096	0.103	0.08
	95% CI	0.456, 0.819	0.577, 0.980	0.517, 0.834
PR	Mean	0.447	0.670	0.475
	SD	0.116	0.184	0.103
	95% CI	0.210, 0.666	0.240, 0.926	0.274, 0.667
QRS	Mean	0.389	0.845	0.371
	SD	0.112	0.198	0.097
	95% CI	0.156, 0.590	0.288, 0.985	0.175, 0.544

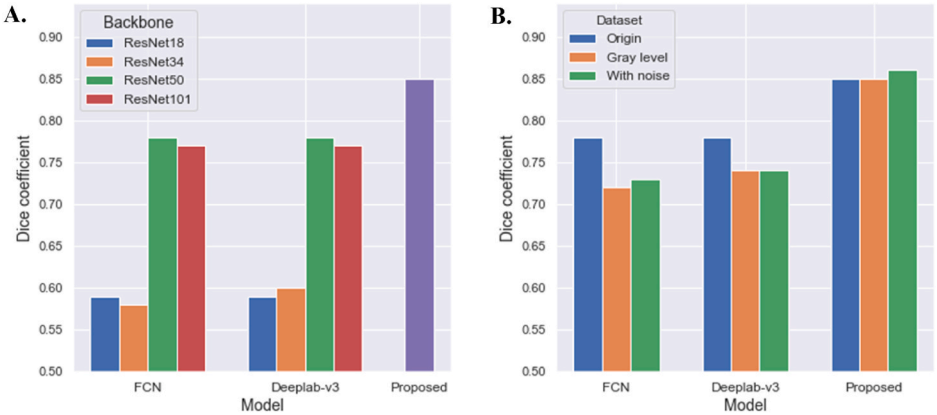

Fig. 5. Segmentation results for the ECG waveform extraction. **A.** Results in cross-validation, where training and testing using different scans of the same type. **B.** Results in the independent testing set, where the models were trained on the binary scan data and tested on another type of ECG image (ResNet50 backbone). In both cases, U-Net achieved the best performance among the three methods.

sufficient to stably measure the parameters. Thus, applying the proposed method to data with longer signals may further improve its performance. Nonetheless, the proposed method could provide good waveform extraction and significant parameter correlation in various layouts.

3.3. The trained model can be applied to another type of ECG image

To showcase the generality of the proposed method, we applied the trained model to an independent testing set with another type of ECG image. We collected 282 coloured ECG images (Supplementary Fig. S3), where the waveform can be labelled by colour thresholding. Then, we degraded them to grey-level to mimic practical situations. In addition, we simulated the image distortion by adding pepper-and-salt noise. The results are summarized in Fig. 5B, and U-Net achieved the best performance. Meanwhile, we observed that all 3 networks achieved acceptable performance, which empirically suggests that a deep learning model trained on low-quality ECG scans can potentially be applied to another type of image. Notably, adding noise slightly increased the performance in our proposed model. We found that the noise weakened both grid and waveform, but the model could denoise the waveform. Thus, adding moderate pepper-and-salt noise may increase the segmentation performance.

4. Discussion

ECG digitization is a classic image-processing problem with increasing attraction due to the eagerness to utilize more old ECG records. However, existing digitization techniques strongly depend on the simple context of scans (e.g., colour thresholding in Ref. [11–13]), which may not capture the more intrinsic geometric property of the ECG waveform. We formulated the digitization problem as a segmentation problem and utilized deep learning. The proposed framework applies U-Net to directly distinguish the waveform from the background, which asks the model to learn the useful features. The model is trained on highly noisy and binary ECG paper records, which forces the model to learn more intrinsic features of the ECG waveform. The results show good segmentation performance and high concordance with the digital signal. Thus, the trained model may also be applied to other types of ECG scans by simply binarizing the image.

Deep learning enables end-to-end formulation for many different problems [32,33] and usually results in better performance than multi-stage methods [34]. experimented on deep learning for ECG image classification without digitization, which challenged that digitization might not have substantial performance gain. Nonetheless, we believe in the necessity of ECG digitization [34]. directly converted the signal to an image and ignored possible noise during printing, storing, and scanning. In addition, digitization enables the reuse of most ECG signal processing methods, eases data sharing and saves storage to enable large-scale studies. Overall, we believe that ECG digitization is essential and can benefit from deep learning.

Arguably, our formulation is relatively straightforward, and U-Net is a general image segmentation network with its own limitations. U-Net does not explicitly capture the strong periodic patterns in ECG scans; thus, integrating the spectrum information into U-Net can be useful. Moreover, U-Net only uses the 3×3 kernel for convolutions and naïvely adds connections between encoder and decoder, which can be improved by using multiple kernel sizes and learnable connections [35]. The overlapping waveform is another problem that cannot be approached with the current segmentation formulation. A potential solution can be lead-by-lead segmentation for a fixed layout. Notably, the direct signal value prediction model may also be trained via methods such as CTC [36], which should ultimately solve the overlapping problem. Nonetheless, we combined the knowledge from previous studies (e.g., Ref. [20,21]) and experimental results to propose a superior architecture for ECG digitization. U-Net can be treated as a strong baseline, which may inspire and guide further development.

5. Conclusions

In this study, we proposed a deep learning framework to digitize paper-based ECG. We formulated the digitization problem as image segmentation and solved it using the U-Net segmentation network. The approach achieved good segmentation and high concordance with the true signal. Since the segmentation model was trained on highly noisy binary images, it can also process ECG scans of a different type in our experiment. We believe that our method can bridge the gap between the application of large-scale real data and the research of advanced ECG analysis algorithms.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to affect the work reported in this paper.

CRediT authorship contribution statement

Yao Li: Conceptualization, Resources, Data curation. **Qixun Qu:** Methodology, Investigation, Software, Formal analysis. **Meng Wang:** Formal analysis. **Liheng Yu:** Resources, Data curation. **Jun Wang:** Conceptualization. **Linghao Shen:** Methodology, Formal analysis, Writing - review & editing, Visualization. **Kunlun He:** Conceptualization, Supervision.

Acknowledgements

This study is supported by the National Key Research and Development Program of China (No. 2018YFC2000701).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2020.104077.

References

[1] B. Vogel, B.E. Claessen, S.V. Arnold, D. Chan, D.J. Cohen, E. Giannitsis, C. M. Gibson, S. Goto, H.A. Katus, M. Kerneis, T. Kimura, V. Kunadian, D.S. Pinto, H. Shiomi, J.A. Spertus, P.G. Steg, R. Mehran, ST-segment elevation myocardial infarction, Nature reviews Disease primers 5 (2019) 39.

[2] G.M. Marcus, Evaluation and management of premature ventricular complexes, Circulation 141 (2020) 1404–1418.

[3] A. Maheshwari, F.L. Norby, N.S. Roetker, E.Z. Soliman, R.J. Koene, M.R. Rooney, W.T. O'Neal, A.M. Shah, B.L. Claggett, S.D. Solomon, A. Alonso, R.F. Gottesman, S. R. Heckbert, L.Y. Chen, Refining prediction of atrial fibrillation-related stroke using the P2-CHA2DS2-VASc score, Circulation 139 (2019) 180–191.

[4] P. Kligfield, L.S. Gettes, J.J. Bailey, R. Childers, B.J. Deal, E.W. Hancock, G. van Herpen, J.A. Kors, P. Macfarlane, D.M. Mirvis, O. Pahlm, P. Rautaharju, G. S. Wagner, M. Josephson, J.W. Mason, P. Okin, B. Surawicz, H. Wellens, American heart association E, arrhythmias committee CoCC, American college of cardiology F and heart rhythm S. Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology a scientific statement from the American heart association electrocardiography and arrhythmias committee, council on clinical cardiology; the American college of cardiology foundation; and the heart rhythm society endorsed by the international society for computerized electrocardiology, J. Am. Coll. Cardiol. 49 (2007) 1109–1127.

[5] T. Reichlin, R. Abacherli, R. Twerenbold, M. Kuhne, B. Schaer, C. Muller, C. Sticherling, S. Osswald, Advanced ECG in 2016: is there more than just a tracing? Swiss Med Wkly 146 (2016) w14303.

[6] S.H. Liu, C.H. Hsieh, W. Chen, T.H. Tan, ECG noise cancellation based on grey spectral noise estimation, Sensors (Basel) 19 (2019).

[7] T. Niederhauser, T. Marisa, L. Kohler, A. Haeberlin, R.A. Wildhaber, R. Abacherli, J. Goette, M. Jacomet, R. Vogel, A baseline wander tracking system for artifact rejection in long-term electrocardiography, IEEE Trans Biomed Circuits Syst 10 (2016) 255–265.

[8] A. Holkeri, A. Eranti, T.V. Kentta, K. Noponen, M.A.E. Haukilahti, T. Seppanen, M. J. Junttila, T. Kerola, H. Rissanen, M. Heliovaara, P. Knekt, A.L. Aro, H.V. Huikuri, Experiences in digitizing and digitally measuring a paper-based ECG archive, J. Electrocardiol. 51 (2018) 74–81.

[9] R.J. Martis, U.R. Acharya, H. Adeli, Current methods in electrocardiogram characterization, Comput. Biol. Med. 48 (2014) 133–149.
[10] G.S. Waits, E.Z. Soliman, Digitizing paper electrocardiograms: status and challenges, J. Electrocardiol. 50 (2017) 123–130.
[11] F. Badilini, T. Erdem, W. Zareba, A.J. Moss, ECG Scan: a method for conversion of paper electrocardiographic printouts to digital electrocardiographic files, J. Electrocardiol. 38 (2005) 310–318.
[12] L. Ravichandran, C. Harless, A.J. Shah, C.A. Wick, J.H. Mcclellan, S. Tridandapani, Novel tool for complete digitization of paper electrocardiography data, JIjoteih and medicine 1 (2013), 1800107-1800107.
[13] M. Baydoun, L. Safatly, O.K.A. Hassan, H. Ghaziri, A. El Hajj, H. Isma'eel, High precision digitization of paper-based ECG records: a step toward machine learning 7 (2019) 1–8. JIJoTEiH and Medicine.
[14] Otsu NJItos, A threshold selection method from gray-level histograms 9 (1979) 62–66, man, and cybernetics.
[15] R.C. Gonzalez, R.E. Woods, B.R. Masters, Digital Image Processing, J. Biomed. Optic. (2009) third ed.
[16] G.S. Waits, Soliman EZJJoe, Digitizing paper electrocardiograms: status and challenges 50 (2017) 123–130.
[17] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015, pp. 234–241.
[18] G. Bradski, Kaehler AJDDsjost, OpenCV 3 (2000).
[19] F. Isensee, J. Petersen, A. Klein, D. Zimmerer, P.F. Jaeger, S. Kohl, J. Wasserthal, G. Koehler, T. Norajitra, S. Wirkert, nnU-Net: Self-Adapting Framework for U-Net-Based Medical Image Segmentation, 2018.
[20] D. Jha, P.H. Smedsrud, M.A. Riegler, D. Johansen, T. De Lange, P. Halvorsen, H. D. Johansen, Resunet++: an advanced architecture for medical image segmentation, in: 2019 IEEE International Symposium on Multimedia (ISM), 2019, pp. 225–2255.
[21] F.I. Diakogiannis, F. Waldner, P. Caccetta, C. Wu, Resunet-a: a deep learning framework for semantic segmentation of remotely sensed data, ISPRS J. Photogrammetry Remote Sens. 162 (2020) 94–114.
[22] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[23] S. Woo, J. Park, J. Lee, I.S. Kweon, CBAM: Convolutional Block Attention Module. European Conference on Computer Vision, 2018, pp. 3–19.
[24] S. Ioffe, CJaL. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, 2015.
[25] A.L. Maas, A.Y. Hannun, A.Y. Ng, Rectifier nonlinearities improve neural network acoustic models, Proc icml 30 (2013) 3.
[26] L. Chen, G. Papandreou, F. Schroff, HJaCV. Adam, P. Recognition, Rethinking Atrous Convolution for Semantic Image Segmentation, 2017.
[27] F. Milletari, N. Navab, S. Ahmadi, V-Net, Fully convolutional neural networks for volumetric medical image segmentation, international conference on 3d vision (2016) 565–571.
[28] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, Focal loss for dense object detection, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 2980–2988.
[29] L. Luo, Y. Xiong, Y. Liu, XJapa Sun, Adaptive Gradient Methods with Dynamic Bound of Learning Rate, 2019.
[30] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, Comput. Vis. Pattern Recogn (2015) 3431–3440.
[31] L.R. Dice, Measures of the amount of ecologic association between species 26 (1945) 297–302.
[32] D. Bahdanau, K. Cho, Bengio YJapa, Neural Machine Translation by Jointly Learning to Align and Translate, 2014.
[33] Y. Taigman, M. Yang, M.A. Ranzato, L. Wolf, Deepface: closing the gap to human-level performance in face verification, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 1701–1708.
[34] R. Brisk, R. Bond, E. Banks, A. Piadlo, D. Finlay, J. McLaughlin, McEneaney DJJoe, Deep learning to automatically interpret images of the electrocardiogram: do we need the raw samples? 57 (2019) S65.
[35] N. Ibtehaz, M.S.J.N.N. Rahman, MultiResUNet : rethinking the U-Net architecture for multimodal biomedical image segmentation 121 (2020) 74–87.
[36] A. Graves, S. Fernández, F. Gomez, J. Schmidhuber, Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks, in: Proceedings of the 23rd International Conference on Machine Learning, 2006, pp. 369–376.