Detection of Myocardial Infarction from 12 Lead ECG Images

Ravikumar Sanjay Sane
Dept. Of Electronics and Electrical Engineering
Indian Institute of Technology Guwahati,
Guwahati, India
s.ravikumar@iitg.ac.in

Pharvesh Salman Choudhary
Dept. Of Electronics and Electrical Engineering
Indian Institute of Technology Guwahati,
Guwahati, India
pharvesh@iitg.ac.in

Dr. L N Sharma
Dept. Of Electronics and Electrical Engineering
Indian Institute of Technology Guwahati,
Guwahati, India
lns@iitg.ac.in

Prof. Samarendra Dandapat
Dept. Of Electronics and Electrical Engineering
Indian Institute of Technology Guwahati,
Guwahati, India
samaren@iitg.ac.in

***Abstract*—Electrocardiogram (ECG) is one of the most practiced methods to detect any heart function abnormalities. ECG results are available as paper records for physical inspection by the medical practitioner. This leads to a large volume of ECG reports resulting in a tedious and error-prone process of examining and retrieving the same. Myocardial Infarction (MI) is the leading cause of the growing global mortality rate and can be detected from ECG. The lack of a physician who is an expert for analysis on the ECG signal is a severe problem, especially in rural areas. This paper proposed a method to detect MI from 12 lead ECG images with the performance of accuracy 87.80%, sensitivity 92.86%, and specificity 78.94%.**

Keywords—Multi-lead ECG printout, scanned images, myocardial infarction(MI), electrocardiogram(ECG), digitization, deep neural network.

I. INTRODUCTION

An electrocardiogram(ECG) is a test that measures the electrical activity of the heart. This electric signal is generated from the human heart to create the cardiac cycle, generating blood circulation. It is composed of three basic components named P wave, QRS complex, and T wave. The P wave is generated during atrial depolarization. After that, the QRS complex is generated when ventricular depolarization occurs, and the T wave is generated during ventricular repolarization.[1] Usually, ECG has to be printed on paper for further physical inspection by a medical practitioner. This leads to a large volume of ECG reports resulting in a tedious and error-prone process of examining and retrieving the same. Further, a lack of a physician who is an expert for ECG signal analysis is one of a severe problem, especially in the rural area.

Myocardial Infarction(MI), also known as a heart attack, is the leading cause of death worldwide and is responsible for 17.3 million death per year. It is expected that this death rate due to heart attack will become 23.6 million by the year 2030. According to the American Heart Association report in 2016, one out of three deaths is a heart attack.[2] Coronary arteries are responsible for the supply of oxygenated blood to the heart. Coronary arteries are blocked by the formation of plaque or blood clots. As a result, the blood supply to the heart decreases, which causes a heart attack, also known as Myocardial Infarction(MI). MI is also known as a silent heart attack because patients are unaware of their condition that they are suffering from MI.[3] An accurate and timely diagnosis of MI is crucial to avoid sudden death. Therefore as discussed above, the lack of expert physicians in rural areas and the cruciality to act stat, there is a need to detect cardiovascular diseases like MI from the multi-lead ECG images.

It is to be noted that a printed multi-lead ECG record includes some characters and annotations, containing information about the patient and the graph itself. While taking the image of such an ECG printout, it is desirable to remove these printed characters. An example of a paper-based ECG printout is shown in figure 1.

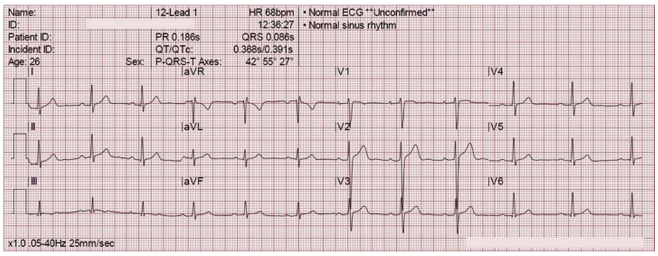

Figure 1. 12 lead ECG image sample[4]

Previous research work is done in the direction of digitization of ECG paper records is referred to in [5-10]. For detection of myocardial infarction from multi-lead ECG is done in [11-14]. [12], [13] and [14] used deep neural networks to classify and localize MI from ECG signals. For ECG classification with SVM with different detail on feature space is referred to in [15-16]. Furthermore wavelet based multiscale energy and eigenspace feature[11] are also used for MI classification from ECG signals.

This paper proposes a method for the detection of myocardial infarction from multi-lead ECG images. The method consists of 2 steps; first, the captured images are de-skewed, and then using various image-processing techniques, the data is retrieved. The second step is to train the deep learning model and detect Myocardial Infarction and Healthy Subject ECG. To validate the data retrieval technique from images and hence measure distortions, the root-mean-square error(RMSE), mean absolute error(MAE) and wavelet PRD are used. To verify the accuracy of the model, this method is compared with the state of the art MEES approach[11]. This paper uses PTB dataset[17] for training and testing the model. The rest of this paper is organized as follows. The proposed data extraction method from ECG images is described in Section II, the proposed deep neural network is described in Section III, and results are produced in Section IV. A brief conclusion is made in Section V.

II. ECG Images to Plot

The method consists of 2 parts, that is, retrieval of data from multi-lead ECG printout and use of this data to detect myocardial infarction using deep neural netwok. This section is dedicated for retrieval of 12 lead ECG data from ECG printout.

A. *ECG Paper Scanning*

The ECG images are scanned with 300 dpi which is equivalent to 295 Hz (3.4 msecs/pixel) sampling rate for data recorded with speed of 25 mm/sec. But scanned images may be skewed. So, de-skewing is done using box-detection algorithm in OpenCV[18] as shown in figure 2. The resultant images might be of different dimensions. So they are resized such that the sampling frequency becomes 250 Hz.

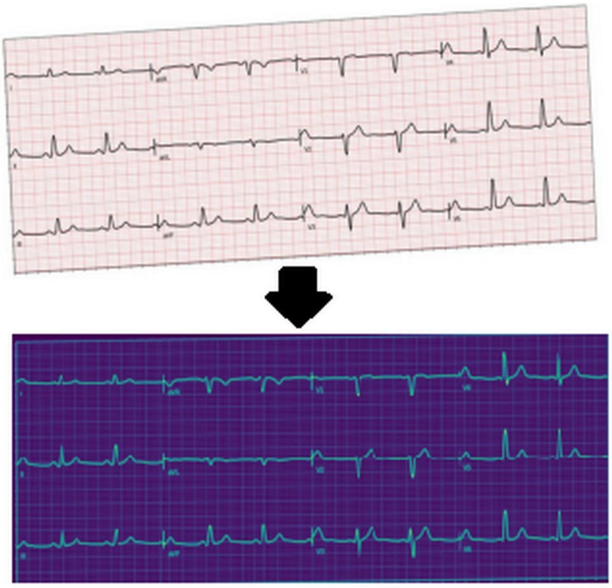

Figure 2. De-skew ECG image

B. *Select the area of interest*

Usually, the standard multi-lead ECG paper consists of 12 leads distributed in 3 rows and 4 columns. Each lead signal is of length 2.5 seconds and covers 12.5 big squares as the recording speed is 25 mm/sec. So in this step, we divide the image into 12 different parts consisting of 3 rows and 4 columns. This step is flexible and depends as per the standard used by the manufacturer of ECG machine. Thus, each 12 leads images are separated as shown in figure 3.

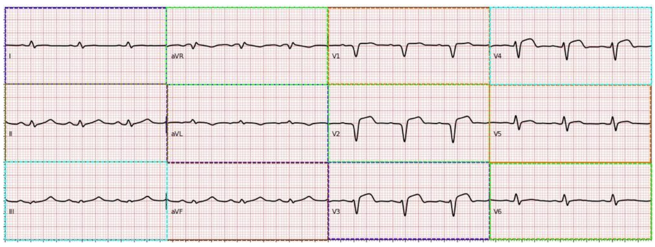

Figure 3. ECG images are separated

C. *RGB to binary image*

This step converts each ECG images into gray scale since the color of ECG signal is black on the paper, while the color of the grid is red. Laplacian filtering is applied for making background noise lighter then the main ECG signal. A threshold of value (125/255) is selected to convert the images to binary form.

D. *Eliminate the annotations from image*

This step uses region based segmentation from the paper[19]. It removes the isolated pixels that do not represent the signal (printed characters). In this step, pytesseract library[20] is also used to detect and crop out any text written on the image.

E. *Binary image to 1D plot*

In this step, a vertical scan is employed to find out the black pixels of thickness more than a threshold (here 3), and has been recorded in an array(figure 4). If no valid pixels are detected for any vertical scan, then the average of previous two values and next 2 values are taken. If the plot goes out of the bound, then only average of previous two values is used.

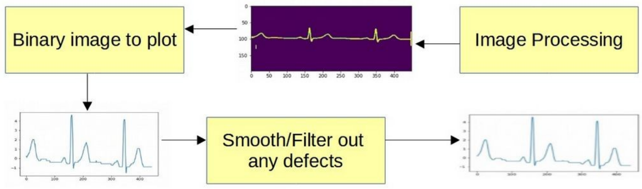

Figure 4. Image to plot conversion

F. *Frequency Adjustment*

During image processing, we have done de-skewing and cropping work, which changes the length and width of image. This changes the number of sampled data points along time-axis(x-axis). To correct this, we use interpolation to correct this defect. Interpolation factor is given by required number of data points divide length of array we got in last step. Here, this paper uses ECG signal sampled at 250 Hz of length 2.5 seconds, so required number of data points is 625.

G. *Pre-processing*

The signal retrieved in the above step is not smooth, and may contain step like changes. This paper hence applies a moving average filter of certain length(here, 5) to smooth out any abnormalities(figure 4).

For detection of myocardial infarction, the DNN uses the shape of the ECG beat. So, this paper uses MinMaxScaler to normalize the retrived signal from the above step to get zero-mean signal. The equation for normalization is shown below.

$$x[n] = \frac{x[n] - \min(x)}{\max(x) - \min(x)}$$

Distortion analysis on the retrieved data is performed and the original data is compared with the retrieved data in order to evaluate the proposed method. Various methods like Root-Mean-Square-Error(rmse), Mean-Absolute-Error(mae) and Wavelet PRD measures are employed to find the fidelity of the data recovered. The figure 5 shows comparison between original and retrieved data of a lead from ECG images.

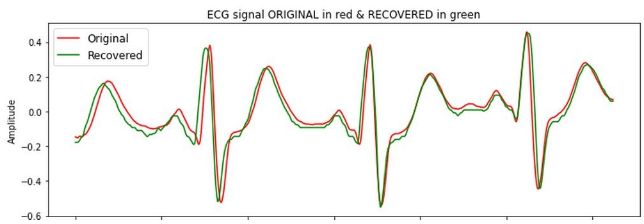

Figure 5. Original vs Retrieved Data of 'I' lead

III. Retrieved Data(Plot) to DNN

In this paper, the PTB database is used to train and test the DNN model. The database contained 549 records from 290 subjects, with 209 males and 81 females. The mean age was 57.2 for males and 61.6 for females. Of the 290 subjects, 148 were labelled as MI with 368 records, 52 were labelled as healthy with 80 records, and the remainder had labels of heart failure, bundle branch block, dysrhythmia, myocardial hypertrophy, valvular heart disease and myocarditis. We divided the PTB database into three classes: "MI", "healthy" and "other" for all other classes. We deleted the records belonging to "other" class and used data from remaining 2 classes, namely, MI and HC.

The PTB database contained 15 simultaneously measured ECG leads (I, II, III, avr, avl, avf, v1, v2, v3, v4, v5 and v6), together with the three Frank lead ECGs (vx, vy and vz). The sampling frequency of the PTB database was set to 1000 Hz. We have not used the Frank lead ECGs in this paper.

As the ECG records were noisy, denoising was performed using an iir notch filter with notch at 50 Hz and a moving average filter with window of size 25. Thereafter, the baseline wander was subtracted from the signal. In this paper, the data is then downsampled to 250 Hz. Thus we have total of 12 leads, with 625 samples(2.5 seconds * 250 Hz) per lead in each frame as input to the network.

In this paper, the state-of-the-art CNN consists of a series of convolutional layers that apply convolutional operations to its inputs in order to detect different spatial patterns [21]: max-pooling layers to provide translational and orientational invariance [21]; dropout layers to provide regularisation and prevent over-fitting [22]; batch normalisation to prevent internal covariate shifting and allow for higher training rates [23]; and fully connected layers to perform the final classification [24]. Table 1 presents the architecture of the CNN classifier used in this study, which was constructed from three convolutional blocks.

Layers	Type	Output Shape
0	Input Layer	(625, 12)
1-3	Conv Block 1	(625, 128)
4-6	Conv Block 2	(625, 256)
7-9	Conv Block 3	(625, 128)
10	Global Avg Pool	(128)
11	Dense	(1)

Table 1. Architecture of CNN classifier

Each convolution block consists of 1d convolution layer, batch normalization layer, and a ReLu activation layer. The last block is followed by a 1d global average pooling layer and a fully connected dense layer.

Type	Filter Size	Kernel/Pool Size
Convolutional 1D	128/256/128	Same
Batch-norm	-	Same
ReLU Act Layer	-	Same

Table 2. Architecture of convolutional block

The processed database is then split into training and testing parts in the 80:20 ratio, as shown in figure 6. Five-fold cross-validation is performed on the classifier trained in this section. For each fold, a separate testing set was used for each fold so that the entire dataset was employed for both training and testing. The training set of the samples were further divided into a 80% training set and 20% validation set as shown in figure 7. The training set was used for back-propagation and weight updating of the neural networks, while the validation set was used for evaluating early stopping and model saving.

The model is trained on the ECG dataset as dicussed above and the data retrived from section II is used to evaluate the model. In order to assess the performance of classifier, the sensitivity, specificity and accuracy for the MI class were computed and compared with the state-of-the-art paper[11].

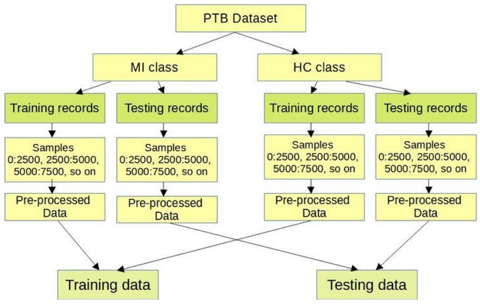

Figure 6. PTB dataset Splits

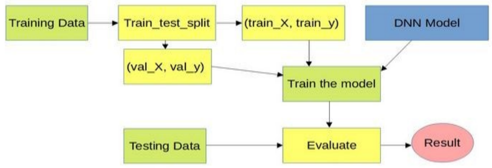

Figure 7. Training the Model

IV. RESULTS

As discussed in section III, the PTB dataset is split into 2 parts as shown in table 3. Then, using testing records, 1927 ECG images are created with the help of modified version ecg_plot library[25] of python.

	Actual Records	Train Records	Test Records	Train Dataset	Test Dataset
MI	148	118	30	5318	1372
HC	54	42	12	1946	555
Total	202	160	42	7264	1927

Table 3. PTB Dataset Split

Then using proposed method from section II, the ECG images were read and the multi-lead data is retrieved back. The retrived data is compared with the original data from which the ECG images were formed. Root-Mean-Square-Error(rmse), Mean-Absolute-Error(mae) and Wavelet PRD measures are calculated to find the fidelity of the data recovered and perform quantitative comparison. The results are attached in table 4. The values of Root-Mean-Square-Error(rmse) obtained is 0.0551 and Mean-Absolute-Error(mae) is 0.0451.

```
    WPRD: (6+1 levels along rows as cA6 cD6 cD5 cD4 cD3 cD2 cD1)*(12 leads along columns)
cA6 0.431,  0.216,  0.392,  0.161,  0.423,  0.279,  0.270,  0.376,  0.400,  0.453,  0.204,  0.308,
cD6 0.235,  0.304,  0.337,  0.306,  0.241,  0.272,  0.371,  0.292,  0.364,  0.488,  0.433,  0.451,
cD5 0.625,  0.396,  0.588,  0.499,  0.561,  0.435,  0.627,  0.386,  0.567,  0.725,  0.521,  0.513,
cD4 0.798,  0.605,  0.864,  0.572,  0.786,  0.793,  0.805,  0.593,  0.708,  0.903,  0.657,  0.659,
cD3 1.170,  1.080,  0.835,  0.793,  0.924,  1.024,  1.206,  0.733,  1.101,  1.257,  0.884,  0.966,
cD2 1.040,  1.615,  1.269,  1.265,  1.151,  1.319,  1.158,  0.944,  1.353,  1.333,  0.997,  1.115,
cD1 0.657,  1.428,  0.707,  0.860,  0.580,  0.718,  0.689,  0.508,  0.783,  2.714,  0.910,  1.046,
```

Table 4. WPRD values of each subbands for 12 ECG leads for 6 level wavelet decompositions

The distortion measured is maximum in cD3, cD2, cD1 bands, which are usually ignored as they contains high frequency components. So it can be concluded that the retrieved data is quite accurate and can be used for detection of Myocardial Infarction.

The deep neural network is trained as descibed in section III and after five-fold cross validation, the confusion matrix, precision, recall and f1-score are calculated with the formula shown below in figure 8. The results are shown in table 5.

sensitivity, recall, hit rate, or true positive rate (TPR)

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{P}} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} = 1 - \mathrm{FNR}$$

specificity, selectivity or true negative rate (TNR)

$$\mathrm{TNR} = \frac{\mathrm{TN}}{\mathrm{N}} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} = 1 - \mathrm{FPR}$$

accuracy (ACC)

$$\mathrm{ACC} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{P}+\mathrm{N}} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}$$

Figure 8. Formula of Accuracy, Recall, Specificity[26]

The accuracy, sensitivity and specificity of the model are calculated in percentage and are compared with the state-of-the-art paper[11]. The paper[11] was implemented and the results are calculated for comparison. The results are shown in table 6.

```
Report & Confusion Matrix
              precision    recall  f1-score   support

          HC       0.79      0.81      0.80       555
          MI       0.92      0.91      0.92      1372

    accuracy                           0.88      1927
   macro avg       0.86      0.86      0.86      1927
weighted avg       0.89      0.88      0.89      1927
```

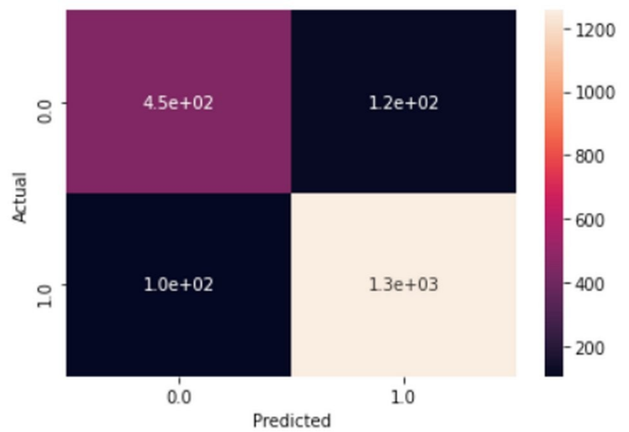

Table 5. Confusion matrix after 5-fold CV

Classifiers	Sensitivity(%)	Accuracy(%)	Specificity(%)
KNN	85	81	77
SVM Lin	90.42	89	87.69
SVM RBF	93	96	99
KNN*	88.23	83.07	69.81
SVM Lin*	81.25	81.73	84.56
SVM RBF*	93.53	91.31	85.77
DNN Model	92.86	87.80	78.94

* (Implemented SOTA paper Result)

Table 6. Performance comparisons with existing method

V. Conclusion

In this paper, we have developed a method to detect myocardial infarction from multi-lead ECG images with the use of deep neaural network. The model achieved the performance of accuracy 87.80%, sensitivity 92.86% and specificity 78.94%. We compared our proposed method with other existing methods. Our model shows comparable results with the existing models which do not use ECG images for MI detection. Further work could explore the training the model on larger dataset like PTB-XL[27] which consists of dataset of 21837 clinical 12-lead ECGs from 18885 patients of 10 second length. The designed deep nerual network model can be further improved and tuned for better performance.

Acknowledgment

The authors are grateful to editor-in-chief and associate ed-itor for encouragement and would like to thank anonymous reviewers for their suggestions to improve the quality of this manuscript.

References

[1] A. C. Guyton and J. E. Hall, “Text Book of Medical Physiology”, 11th ed.New York, NY, USA: Elsevier, 2006.

[2] X. Wang and K. J. A. Peter, "Mending the failing heart," vol. 11, no. 6, p. 1605, 2019.

[3] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. J. I. S. Adam, “Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals”, vol. 415, pp. 190-198, 2017

[4] http://www.ems12lead.com/wp-content/uploads/sites/42/2010/01/normal-ecg.png

[5] Tsair Kao, Len-Jon hwang, Yui Han Lin, Tzong huei Lin, Chia Hung Hsio, “Computer Analysis of Electrocardiograms from ECG paper Recordings” Proceeding of 23rd annual EMBS international conferences, Istanbul, Turkey, 2001.

[6] T W Shen, T F Laio Image, “Processing on ECG Chart for ECG Signal Recovery”, Tzu Chi University, Hualien, Taiwan Computers in Cardiology 2009; 36:725728

[7] Ricardo Baeza, “Electrocardiogram digitization: a practical perspective on the usefulness of a new tool to convert paper electrocardiograms into digital waveform”, Journal of Electrocardiology, vol. 38, issue 4, pp. 321-323, October 2005

[8] Fabio Badilini, Tanju Erden, Wojciech Zareba, and Arthur J. Moss, “ECG scan: a method for conversion of paper electrocardiographic printouts to digital electrocardiographic files”, Journal of Electrocardiology, vol.38, issue 4, pp. 310-318, October 2005.

[9] Prashanth Swamy, Srinivasan Jayaraman, and M. Girish Chandra, “An Improved Method for Digital Time Series Signal Generation from Scanned ECG Records”, International Conference on Bioinformatics and Biomedical Technology, pp. 400-403, 2010.

[10] Shuang Wang, Shugang Zhang, Zhen Li, Lei Huang, Zhiqiang Wei, “Automatic digital ECG signal extraction and normal QRS recognition from real scene ECG images”, Elsevier, Nov 2019

[11] L. N. Sharma, R. K. Tripathy, and S. Dandapat, “Multiscale Energy and Eigenspace Approach to Detection and Localization of Myocardial Infarction”, IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 62, NO. 7, JULY 2015.

[12] J.F. Wu, Y.L. Bao, S.C. Chan, H.C. Wu, L. Zhang, X.G. Wei, “Myocardial infarction detection and classification-A new multi-scale deep feature learning approach”, Int. Conf. Digit. Signal Process. DSP (2017), pp. 309-313.

[13] U.R. Acharya, H. Fujita, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, “Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals”, Inf Sci, 415–416 (2017), pp. 190-198.

[14] Lui, Chow, H.W. Lui, K.L. Chow, “Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ECG devices”, Informatics in Medicine Unlocked, 13 (2018), pp. 26-33.

[15] Qibin Zhao and Liqing Zhang, “ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines”, Neural Networks and Brain, 2005(ICNN&B '05), pp. 1089-1892, 2005.

[16] Nasiri J.A., Naghibzadeh M., Yazdi, H.S. and Naghibzadeh B., “ECG Arrhythmia Classification with Support Vector Machines and Genetic Algorithm”, Computer Modeling and Simulation, (EMS '09), pp. 187-192, 2009.

[17] https://www.physionet.org/content/ptbdb/1.0.0/ PTB Diagnostic ECG Database, Published: Sept. 25, 2004. Version: 1.0.0.

[18] https://docs.opencv.org/master/dd/d49/tutorial_py_contour_features.html

[19] Deepak Kumar Garg, Diksha Thakur, Seema Sharma, Shweta Bhardwaj, “ECG Paper Records Digitization throughImage Processing Techniques”, International Journal of Computer Applications (0975 –888), Volume 48–No.13, June 2012.

[20] https://tesseract-ocr.github.io/

[21] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. “Recent advances in convolutional neural networks.” ArXiv 2015:1–14.

[22] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting”, J Mach Learn Res, 15 (2014), pp. 1929-1958.

[23] M. Windows, Os, “Batch normalization: accelerating deep network training by reducing internal covariate shift”, Uma Ética Para Quantos, XXXIII (2014), pp. 81-87.

[24] Y. Lecun, Y. Bengio, G. Hinton, "Deep learning", Nature, 521 (2015), pp. 436-444.

[25] https://pypi.org/project/ecg-plot/ ECG Author: dy1901@gmail.com

[26] https://en.wikipedia.org/wiki/Confusion_matrix

[27] https://www.physionet.org/content/ptb-xl/1.0.1/ Published: April 24, 2020. Version: 1.0.1