TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Phân lớp ảnh điện tâm đồ bằng kỹ thuật học sâu hỗ trợ chẩn đoán bệnh tim mạch

Lê Xuân Vinh[1,*], Lê Xuân Việt[1], Vũ Hoàng Thương[2]

[1]Khoa Công nghệ Thông tin, Trường Đại học Quy Nhơn, Việt Nam
[2]Liên hiệp hội Khoa học Kỹ thuật Bình Định, Việt Nam

Ngày nhận bài: 26/02/2021; Ngày nhận đăng: 12/05/2021

TÓM TẮT

Trong bài báo này, chúng tôi đề xuất mô hình phân lớp ảnh điện tâm đồ để hỗ trợ cho các chuyên gia y tế trong việc chẩn đoán bệnh tim mạch. Mô hình được xây dựng bằng kỹ thuật học sâu, được huấn luyện và kiểm tra bằng dữ liệu MIT-BIH. Cụ thể hơn, mạng nơ-ron tích chập được thiết kế với đầu vào là ảnh lấy từ điện tâm đồ và đầu ra được gắn nhãn là loại rối loạn nhịp tim tương ứng nếu có. Mô hình thu được đáng tin cậy với kết quả thử nghiệm có độ chính xác cao từ 97 - 99%.

Từ khóa: *Phân lớp, học sâu, mô hình, mạng noron tích chập, điện tâm đồ.*

**Tác giả liên hệ chính.*
Email: lexuanvinh@qnu.edu.vn

https://doi.org/10.52111/qnjs.2021.15505

JOURNAL OF SCIENCE
QUY NHON UNIVERSITY

Classification of ECG images by Deep Learning to support the diagnosis of cardiovascular diseases

Le Xuan Vinh[1,*], Le Xuan Viet[1], Vu Hoang Thuong[2]

[1]Faculty of Information Technology, Quy Nhon University, Vietnam
[2]Binh Dinh Union of Science and Technology Associations, Vietnam

Received: 26/02/2021; Accepted: 12/05/2021

ABSTRACT

In this paper, we have proposed the electrocardiogram (ECG) classification model to partly support medical professionals. The model is built by deep learning technique with MIT-BIH training data. Specifically, the convolutional neural network (CNN) will be built to take input of images cropped from the ECG images and the output labeled as the corresponding arrhythmia. The obtained model is reliable when the test results have a relatively high accuracy of 97 to 99%.

Keywords: *Classification, deep learning, model, convolution neural network, electrocardiogram.*

1. INTRODUCTION

The electrocardiogram (ECG) is a chart that records changes in electrical current in a person's heart over a given period of time. The heart contracts itself in rhythm thanks to the heart muscle's control of its conduction system. The difference in concentration of Kali, Canxi ions on the two sides of the membrane creates a potential. The electrical current of the heart is as small as a few parts per a thousand volts. But the current is amplified, and its signal is plotted on the paper which is called ECG.

Until now, the ECG is a very popular method of monitoring the electrical activity of the heart. Most clinics, hospitals use the ECG to aid in the diagnosis of heart rhythm problems and structural abnormalities. An ECG analyst can diagnose a patient's cardiovascular disease, which requires many experiences and times from doctors. Computer-oriented approaches can effectively reduce diagnostic time, increase the number of ECG records processed, support to local health facilities where there are not many experienced ECG analysts.

According to that approach, there have been many different studies. Some classification results of ECG signals with typical methods such as frequency analysis,[1] statistical methods,[2] heuristic-based methods,[3] artificial neural networks,[4] support vector machines (SVM),[5] wavelet transform,[6] and hidden Markov models.[7] Artificial neural networks obtained an average accuracy of 90.6% for the classifier of ECG wave into six classes[8] and a feed-forward neural networks was used as a classifier for the detection of four types of arrhythmia classes

*Corresponding author.
Email: lexuanvinh@qnu.edu.vn

and obtained an average accuracy of 96.95%.[9] Most of these results are obtained by using the traditional approach to solving the classification problems on ECG data.

Machine learning has recently developed and become an approach that is able to solve the classification issue effectively in medicine. Deep learning, a subset of machine learning, has a wide range of applications in the prediction and prevention of fatal sicknesses, particularly cardiovascular diseases. A recurrent neural network achieved an average accuracy of 98.06% for detecting four types of arrhythmia.[10] A convolutional neural network model proposed for the classification from a 1-D ECG signal[11] yielded a classification accuracy of 96.72%. Another deeper 1-D CNN model proposed for the classification of ECG dataset[12] obtained an average accuracy of 97.3%. A nine-layer 2-D CNN model proposed for the classification of five different heartbeat arrhythmia types[13] achieved an accuracy of 94.03%.

As mentioned above, building the ECG classification model is an urgent need in practice, which is posed by local medical administrators. The model will be a key component of the software assisting ECG analysts. Therefore, we build the model of ECG signal image classification using deep learning technique with MIT-BIH[14] training data. Specifically, the convolutional neural network (CNN) will be built to take input as cropped images of the ECG image, and the output is labeled as the corresponding arrhythmia.

This paper is organized as follows. Section 2 presents the features of ECG data, image data obtained from the MIT-BIH database. Section 3 then proposes a CNN structure with its components and training methods to build a modeling of ECG classification. Section 4 discusses the training for the model and its accuracy. Finally, the conclusions and comments are in Section 5.

2. CHARACTERISTICS OF ECG DATA

The original data from the MIT-BIH database is one-dimensional data. From this database, the ECG images were reproduced and published on the Kaggle.[15] The ECG images of each type of arhythmia were labeled with the name of the directory in which it resides. We use this image dataset to train and evaluate the model.

Every image is a cycle of ECG. Its characteristics are waves criteria (P, Q, R, S, T, U) and the time interval between them. Several normal and typical ECG cycles are plotted as shown in the Figure 1 below.

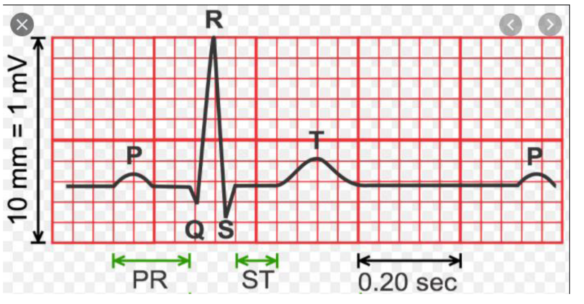

Figure 1. One cycle of ECG

The database includes 15 characteristic morphological types divided into 5 classes: (i) normal beat (N), (ii) supraventricular ectopic beat (S), (iii) ventricular ectopic beat (V), (iv) fusion beat (F), and (v) paced beats or unknown beat (Q). Details of each class are as follows:

(i) Class N consists of

R - Right bundle branch block beat

N - Normal beat

L - Left bundle branch block beat

j - Nodal (junctional) escape beat

e - Atrial escape beat

(ii) Class S consists of

S - Supraventricular premature beat

J - Nodal (junctional) premature beat

a - Aberrated atrial premature beat

A - Atrial premature beat

(iii) Class V consists of

V - Premature ventricular contraction

E - Ventricular escape beat

(iv) Class F consists of

F - Fusion of the ventricular and normal beat

(v) Class Q consists of

Q - Unclassifiable beat

P - Paced beat

f - Fusion of the paced and normal beat

The data are divided into two sets: training and testing. The number of images in each episode is as follows.

Table 1. Number of images in the datasets

Set	F	N	Q	S	V
Training	641	22122	6413	2222	5780
Testing	161	13661	1608	557	1448

Now, we will build a CNN network to classify the dataset just presented above.

3. THE ECG IMAGES CLASSIFIER

We know that ECG signal data has many different representations. There is a suitable treatment for each respective representation. Image data format as shown in the above section, suitable for the classification model using Convolution neural network (CNN) because the CNN has the ability to automatically extract image feature. CNN is a development of a deep neural network (DNN). Its number of parameters is less than that of DNN because it uses share weights. Thanks to this property, local features on the image are extracted, and there is no need to match the entire image, which is consistent with the ECG image's feature when processed.

The structure of CNN was first proposed by Le Cun et al. in 1989[16] and then it was improved.[17] Then, they developed a multi-layer artificial neural network called LeNet-5 which can be trained with the backpropagation algorithm[18]. Since 2006, many methods have been developed to overcome the difficulties encountered in training deep CNNs.[19-21] Most notably, AlexNet[22] is similar to LeNet-5 but with a deeper structure. Then, many studies have been proposed to improve its performance. Among them are four representations studies of ZFNet,[22] VGGNet,[23] GoogleNet[21] and ResNet.[24]

The classifier is built from the VGG16 network architecture. Basically, it consists of three types of layers, namely convolution, pooling, and fully connected layers.

Table 2. Architecture of the classifier

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(128, 128, 3)]	0
block1_conv1 (Conv2D)	(128, 128, 64)	1792
block1_conv2 (Conv2D)	(128, 128, 64)	36928
block1_pool (MaxPooling2D)	(64, 64, 64)	0
block2_conv1 (Conv2D)	(64, 64, 128)	73856
block2_conv2 (Conv2D)	(64, 64, 128)	147584
block2_pool (MaxPooling2D)	(32, 32, 128)	0
block3_conv1 (Conv2D)	(32, 32, 256)	295168
block3_conv2 (Conv2D)	(32, 32, 256)	590080
block3_conv3 (Conv2D)	(32, 32, 256)	590080
block3_pool (MaxPooling2D)	(16, 16, 256)	0
block4_conv1 (Conv2D)	(16, 16, 512)	1180160
block4_conv2 (Conv2D)	(16, 16, 512)	2359808
block4_conv3 (Conv2D)	(16, 16, 512)	2359808
block4_pool (MaxPooling2D)	(8, 8, 512)	0
block5_conv1 (Conv2D)	(8, 8, 512)	2359808
block5_conv2 (Conv2D)	(8, 8, 512)	2359808
block5_conv3 (Conv2D)	(8, 8, 512)	2359808
block5_pool (MaxPooling2D)	(4, 4, 512)	0
flatten (Flatten)	(8192)	0
fc1 (Dense)	(4096)	33558528
dropout (Dropout)	(4096)	0
fc2 (Dense)	(4096)	16781312
dropout_1 (Dropout)	(4096)	0
predictions (Dense)	(5)	20485

Total params: 65,075,013

Trainable params: 50,360,325

Non-trainable params: 14,714,688

The convolutional layer aims to learn feature of the inputs, including some of the kernels which are used to compute different feature maps. Specifically, each neuron of a feature map is connected to a region of neighbouring neurons in the previous layer. Such a neighbourhood is referred to as the neuron's receptive field in the previous layer. Note that, to generate each feature map, the kernel is shared by all spatial locations of the input. The built CNN contains 5 convolutional blocks. In which, the first 2 blocks, each block has 2 convolution layers and the remaining 3 blocks each block has 3 convolution layers.

The pooling layers are used to reduce the dimensions of the feature maps. Thus, it reduces the number of parameters to learn and the amount of computation performed. It is usually placed between two convolutional layers. Thus, there are five pooling layers and use 2D max-pooling operator.

Finally, two fully connected layers use the Relu function as activation function and Dropout to avoid overfitting. The output layer which has 5 nodes corresponding to 5 classes and softmax function is used here. The model structure is summarized in Table 2.

The classes of this model can be seen in more detail in the similar model of A. Ullah [29]

The data for the input layer are images (128, 128, 3) converted from PNG images (256, 256, 3) in the training dataset by cv2.resize( ).

4. EXPERIMENTS AND RESULTS

The work of evaluating the accuracy of the model involves training the classifier. So in the first part of this section, we will cover that process.

Recall that the data is used from the two training and testing sets as mentioned in Section 2. To make an easier and simpler machine learning model, in this paper we have converted to numerical data in npy style files. The testing data are taken as the validation data during training.

The loss function is a measure of how good the neural network is during training and is represented by the difference between a given training pattern and the network outputs. There are many different types of loss functions, but deep learning often uses cross-entropy function, which is

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left(y_i \ln(a_i) + (1-y_i)\ln(1-a_i)\right) \quad (1)$$

Where N is the number of training data or the batch size, y_i is 0,1 depending on the class label and a_i is an actual value from the output layer. In case of multiple classes, the loss function L is extended from Formula 1.

To minimize the loss function, there are several well-known optimization algorithms such as Adam[25], Adagrad[26], and Adadelta[27]. Here, we choose the Adam algorithm in the training process.

The role of the activation function is to determine the output value of each node based on the input signal composite value. The nonlinear activation is widely used, including rectified linear units (ReLU), leakage rectified linear units (LReLU), and exponential linear units (ELU). Among them, the ReLU function is the most widely used, and is used in this model.

Our programs were implemented in Python programming language, using the framework of Keras. The training environment used is Google Colab. It takes about 0.5 - 1 hours to get results after 100 epochsdepending on the GPU provided by Google. The training of deep learning model could early stop scheme by using validation set. However, to gain experience, we have tried to train up to 200 epochs and found that just 100 epochs is enough to achieve a good model.

Another problem is the dropout during training. Dropout is first introduced by Hinton et al. [28], and it has been proven to be very effective in reducing overfitting. We apply Dropout to the fully connected layers. The dropout with the parameter 0.3 produces a model with good training accuracy (99.07%), but poor validation accuracy (96.05%). Meanwhile, a dropout with parameter 0.5 results in a model with poor training accuracy (96.95%), but better validation accuracy (97.52%), perhaps avoid overfitting.

With parameters such as the batch-size is 64, dropout is 0.4, we run through 100 epochs on the training data set, the validation data is also the testing data and model accuracy and loss for the model is shown in Figure 2.

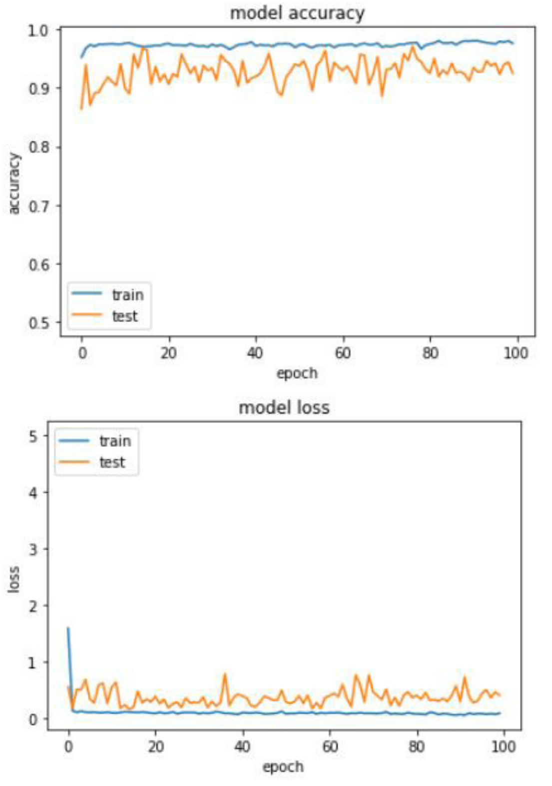

Figure 2. The model accuracy and loss

Performance of the classifier has an accuracy of 97.57% on the testing data set and an accuracy of 99.13% on the training data set.

The confusion matrix of classifier on testing data set is presented in Figure 3. Performance of the classifier has an accuracy of 97.57% .

In Figure 3, the labels of the classes corresponding to 0, 1, 2, 3, 4 are F, N, Q, S, V.

Observing the confusion matrix in Figure 3, we see that the classification performance of the model is of good quality, especially for the S and V classes. Clinically, supraventricular ectopic beats (S) and ventricular ectopic beats (V) are two critically abnormal and serious heartbeats, and the performance of the model in the test for S and V classes is very high.

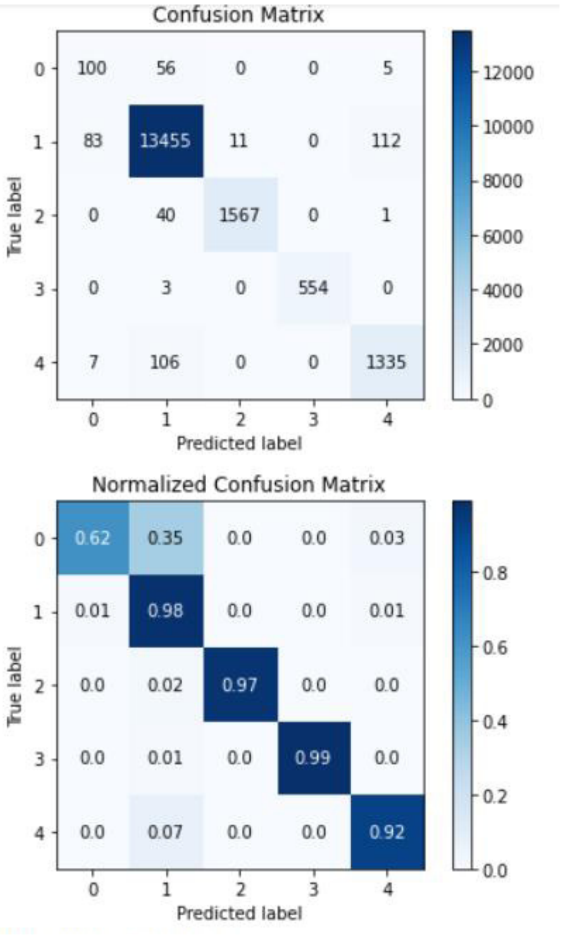

Figure 3. Confusion Matrix for testing data

Training data are also collected from patients, and we get it from the MIT-BIH database. We present more predictive results of the model on the training data set in Figure 4 for your reference. Of course, the model's classification performance on this data set is higher. Performance of the classifier has an accuracy of 99.13%.

5. CONCLUSIONS

In this paper, we have proposed the building of an ECG image classifier using deep learning techniques. The model is obtained by training a 2D-CNN on a preprocessing image set from the MIT-BIH database which has been evaluated on test data and obtained a quite high accuracy, especially the classification of abnormal and serious heartbeats. More importantly, we propose a method to build the ECG image classifier including the CNN structure and the training and testing procedure for the model when we have a new data set. This classification model will be an important component in the ECG analysis support software for physicians, medical professionals diagnosing cardiovascular disease, an urgent requirement of local health facilities.

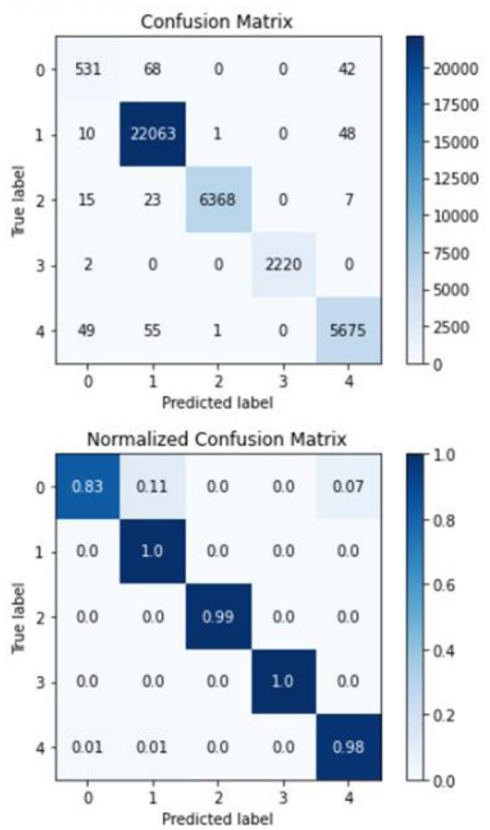

Figure 4. Confusion Matrix for training data

REFERENCES

1. K. I. Minami, H. Nakajima, T. Toyoshima. Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network, *IEEE Transaction on Bio-medical Engineering,* **1999**, 46, 179–185.

2. J. L. Willems, E. Lesaffre. Comparison of multigroup logistic and linear discriminant ecg and vcg classification, *Journal of Electrocardiology,* **1987**, 20, 83–92.

3. S. Osowski, L.T. Hoai, T. Markiewicz. Support vector machine based expert system for reliable heartbeat recognition, *IEEE Transaction on Bio-medical Engineering,* **2004**, 51, 582–589

4. R. Ceylan, Y. Ozbay. Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network. *Expert System and Applications*, **2007**, 33, 286–295.

5. A. Mustaqeem, S. M. Anwar, M. Majid. Multiclass classification of cardiac arrhythmia using improved feature selection and SVM invariants, *Computatinal and Mathematical Methods in Medicine,* **2018,** 1**,** 1-10.

6. O. T. Inan, L. Giovangrandi, G. T. Kovacs. Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features, *IEEE Transaction on Bio-medical Engineering,* **2006**, 53, 2507–2515.

7. D. A. Coast, R. M. M. Stern, G. G. Cano, S. A. Briller, An approach to cardiac arrhythmia analysis using hidden markov models, *IEEE Transaction on Bio-medical Engineering,* **1990**, 37, 826–836.

8. L. Dehan, X. U. Guanggui, Z. Yuhua, H. G. Hosseini. Novel ECG diagnosis model based on multi-stage artificial neural networks, *Chinese Journal Science Instruments,* **2008**, 29, 27-38.

9. R. Ceylan, Y. Ozbay. Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network. *Expert System and Applications,* **2007**, 33, 286–295.

10. C. Salvatore, A. Cerasa, P. Battista, M. C. Gilardi, A. Quattrone, I. Castiglioni. Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: A machine learning approach, *Frontiers in Neuroscience,* **2015**, 9, 307-319.

11. S. Kiranyaz, T. Ince, M. Gabbouj. Real-time patient-specific ECG classification by 1-D convolutional neural networks, *IEEE Transaction on Bio-medical Engineering*, **2015**, 63, 664–675.

12. P. Rajpurkar, A. Y. Hannun, Haghpanahi. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, *Nature Medicine,* **2019**, 25, 65-69.

13. U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, R. San Tan. A deep convolutional neural network model to classify heartbeats, *Computatinal and Mathematical Methods in Medicine,* **2017**, 89, 389–396.

14. G. B. Moody and R. G. Mark. The impact of the MIT-BIH arrhythmia database, *IEEE Engineering in Medicine and Biology Magazine*, **2001**, 20 (3), 45–50.

15. Kaggle, <https://www.kaggle.com/analiviafr/ecg-images?select=ecg_img>, Retrieved January 27, 2021.

16. B. B. Le Cun, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, *Handwritten digit recognition with a back-propagation network*, Proceedings of the Advances in Neural Information Processing Systems (NIPS), 1989, 396–404.

17. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner. *Gradient-based learning applied to document recognition*, in: Proceedings of IEEE, 1998, 2278–2324.

18. R. Hecht-Nielsen. Theory of the backpropagation neural network, *Neural Networks 1* (Supplement-1), **1988**, 445–448.

19. X.-X. Niu, C. Y. Suen. A novel hybrid cnn–svm classifier for recognizing handwritten digits, *Pattern Recognition,* **2012,** *45*(4), 1318–1325.

20. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. Imagenet large scale visual recognition challenge, *International Journal of Conflict and Violence* (IJCV), **2015**, *115*(3), 211–252.

21. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich. *Going deeper with convolutions*, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, 1–9.

22. M. D. Zeiler, R. Fergus. *Visualizing and understanding convolutional networks*, in: Proceedings of the European Conference on Computer Vision (ECCV), 2014, 818–833.

23. O. Russakovsky, J. Deng. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision,* **2015,** 115, 21-252.

24. K. He, X. Zhang, S. Ren, J. Sun. *Deep residual learning for image recognition*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, 770–778.

25. A. Emelogu, S. Chowdhury, M. Marufuzzaman. An enhanced sample average approximation method for stochastic optimization, *International Journal of Production Economics,* **2016**, 182, 230-252.

26. J. Duchi, E. Hazan, Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization, *Journal of Machine Learning Research*, **2011**, 12, 2121-2159.

27. R. Zhang, W. Gong, V. Grzeda, A. Yaworski. An adaptive learning rate method for improving adaptability of background models, *IEEE Signal Processing Letters*, **2013**, 20, 1266–1269.

28. K. Nguyen, C. Fookes. *Improving deep convolutional neural networks with unsupervised feature learning*, in: proceding of IEEE International Conference on Image Processing, 2015, 2270-2274.

29. A. Ullah, S.M. Anwar, M. Bilal, R.M. Mehmood. Classification of Arrhythmia by Using Deep Learning with 2-D ECG Spectral Image Representation, *Remote Sensing,* **2020**, *12*(10), 1685-1699.