# ECG CLASSIFICATION USING HIGHER ORDER SPECTRAL ESTIMATION AND DEEP LEARNING TECHNIQUES

*H. Alquran*[*], *A.M. Alqudah*[*], *I. Abu-Qasmieh*[*], *A. Al-Badarneh*[*], *S. Almashaqbeh*[*]

**Abstract:** Electrocardiogram (ECG) is one of the most important and effective tools in clinical routine to assess the cardiac arrhythmias. In this research higher-order spectral estimations, bispectrum and third-order cumulants, are evaluated, saved, and pre-trained using convolutional neural networks (CNN) algorithm. CNN is transferred in this study to carry out automatic ECG arrhythmia diagnostics after employing the higher-order spectral algorithms. Transfer learning strategies are applied on pre-trained convolutional neural network, namely AlexNet and GoogleNet, to carry out the final classification. Five different arrhythmias of ECG waveform are chosen from the MIT-BIH arrhythmia database to evaluate the proposed approach. The main contribution of this study is to utilize the pre-trained convolutional neural networks with a combination of higher-order spectral estimations of arrhythmias ECG signal to implement a reliable and applicable deep learning classification technique. The Highest average accuracy obtained is 97.8 % when using third cumulants and GoogleNet. As is evident from these results, the proposed approach is an efficient automatic cardiac arrhythmia classification method and provides a reliable recognition system based on well-established CNN architectures instead of training a deep CNN from scratch.

Key words: *bispectrum, third cumulants, transfer learning, deep learning, convolutional neural network*

*Received: December 20, 2018* **DOI:** 10.14311/NNW.2019.29.014
*Revised and accepted: August 27, 2019*

## 1. Introduction

The electrocardiogram (ECG) examines the electrical activity of the heart and cardiac health. Using ECG, physicians can detect heart disorders such as heart arrhythmias, cardiac abnormalities, and heart attack. Healthy subjects have a certain ECG morphological shape, therefore, any changing on the distinctive ECG signal shape indicates an irregularity of heart rhythm, or even damaging of the

[*]Hiam Alquran; Ali Mohammad Alqudah – Corresponding author; Isam Abu-Qasmieh; Alaa Al-Badarneh; Sami Almashaqbeh; Department of Biomedical Systems and Informatics Engineering, Yarmouk University, Irbid, Jordan, E-mail: heyam.q@yu.edu.jo, ali_qudah@hotmail.com, ali_qudah@yu.edu.jo

 

heart muscle. In accordance to the Association for the Advancement of Medical Instrumentation (AAMI) categorization, the cardiac arrhythmias are divided into 5 major classes; Normal (N), Supraventricular ectopic (S), Ventricular ectopic (V), Fusion (F) and Unknown (Q) [1–3]. Analysis, Detection, and classification of different ECG waveforms are crucial issues in cardiology clinical practice. These stages are time-consuming, and they are highly subjective to errors [1, 2]. Machine learning techniques are employed to analyze, detect, and classify the different ECG waveforms, which makes the diagnostic decision more precise and treatment plan more efficient.

Because of the diversity of ECG signals among subjects, various researches were carried out to study and to classify them and consequently, several algorithms have been developed and applied to classify ECG signal [4, 5]. Acharya et al. [1] presented a 9-layer deep Convolutional Neural Network (CNN) to classify the five heartbeat types. The average accuracy of their approach was 93.5 %. Shen et al. extracted wavelet-based features with other features (like Morphological, spectral and statistic) to detect a cardiac arrhythmia in the ECG, where they used Support Vector Machine (SVM) as a classifier with an average recognition rate of 98.9 %.

Also, the research study of Joshi and Topannavar [6] used SVM as the five heartbeat types classifier, where two features, morphological (wavelet transform) and dynamic (RR interval), were extracted from the ECG signal, and the accuracy of their classifier in the class-oriented evaluation was 99 %, and 86 % in the subject-oriented evaluation. Firoozabadi et al. [7] focused on his research on employing linear Discriminant Function Analysis (DFA) to identify ischemia caused by exercise during an exercise stress test, where Morphological features (QRS slopes) were extracted with Sensitivity of 24 % and specificity of 93 %.

Banerjee and Mitra [2] proposed a method for analyzing ECG patterns using the cross wavelet transform (XWT) and Independent Component Analysis (ICA). They used a threshold-based classifier to classify the ECG signal into two classes: normal class and Inferior Myocardial Infarction (IMI) class with an overall accuracy of 97.6 %. Venkatesan et al. applied time domain and frequency domain features in KNN classifier to classify the ECG signals into two classes, where the accuracy was 97.5 %.

This paper represents a new classification approach for extracting spectral features from the ECG signal in order to recognize the five ECG arrhythmias, Normal, Left bundle branch block (LBB), Right bundle branch block (RBB) form N category, Premature ventricular contraction (PVC) from V category, and Atrial premature beat (APB) from S category, by applying higher-order spectral estimation and transfer learning strategies on pre-trained CNN.

## 2. Materials and methods

The proposed methodology in this paper mainly consists of three main parts: (1) ECG preprocessing and beat segmentation, (2) Calculating Higher-Order Spectral Feature (Bispectrum and Third Order Cumulants), and (3) classification of ECG beat using pretrained CNN (GoogleNet and AlexNet). Finally, a comparison is made between the proposed techniques. The block diagram of the proposed methodology is shown in Fig. 1.

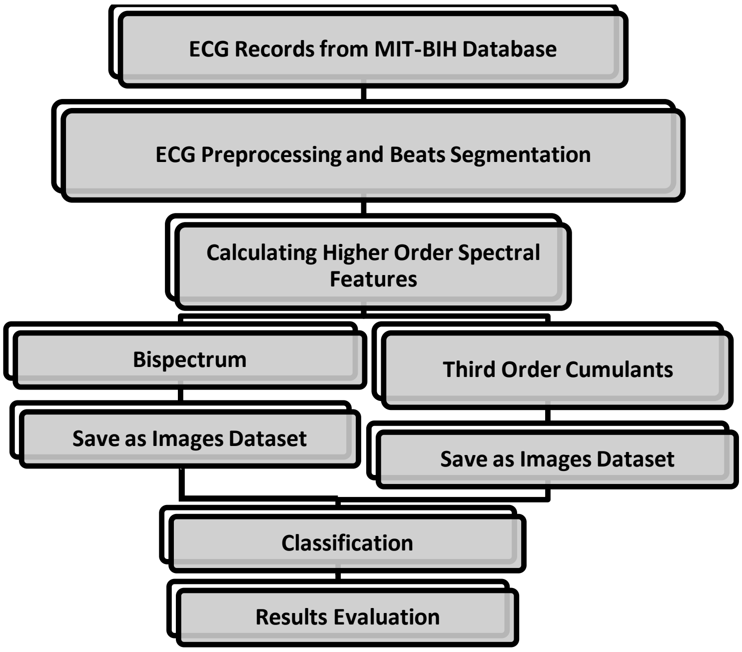


**Fig. 1** *Block diagram of the proposed method.*

## 2.1 ECG database

This paper has utilized one of the most commonly used ECG databases, which is the MIT-BIH arrhythmia database. The database is divided into training and testing sets for evaluating and validating the proposed classifier methodology using higher-order spectral estimation. MIT-BIH arrhythmia database consists of 48 ECG recordings, each one of 30 min in length segment selected from 24-hour recordings of 48 patients [8,9]. In this paper, the normal and arrhythmic beats are annotated based on AAMI standard, and the present methodology is designed to classify the beats into five classes: normal (N), left bundle branch block (LBBB), right bundle branch block (RBBB), premature ventricular contraction (PVC), and atrial premature beat (APB) [9]. The distribution of beat arrhythmia types of the selected database is listed in Tab. I.

| Beat Type | Record Number |
|---|---|
| Normal | 103, 113, 115, 123, 220, 234 |
| LBBB | 109, 111, 207, 214 |
| RBBB | 118, 124, 212, 231 |
| APB | 209, 222, 232 |
| PVC | 106, 116, 119, 200, 203, 208, 210, 213, 215, 221, 233 |

**Tab. I** *Distribution of ECG records of five different ECG beats.*

## 2.2 Pre-processing ECG beat and segmentation

In this section, the pre-processing techniques used for filtering the ECG records and the algorithm of ECG beats segmentation are described. The preprocessing techniques used for filtration of ECG signal consists of the following steps; applying Butterworth bandpass filter with a frequency range of 0.1 and 100 Hz to eliminate the baseline wandering and the high-frequency noise. Next, the smoothing of the ECG signal was done by applying a moving average filter [$8-12$]. The moving average filter is one of the most widely used noise reduction filters in signal processing and especially biomedical signal processing. In this type of filter, an overlapped window with predefined size moves over the signal, and for each window, the average for the values inside the window is calculated and inserted at the window center, which results in the reduction of the extreme or outlier values that represent noise. Fig. 2 shows the ECG signal beat before and after the preprocessing techniques.

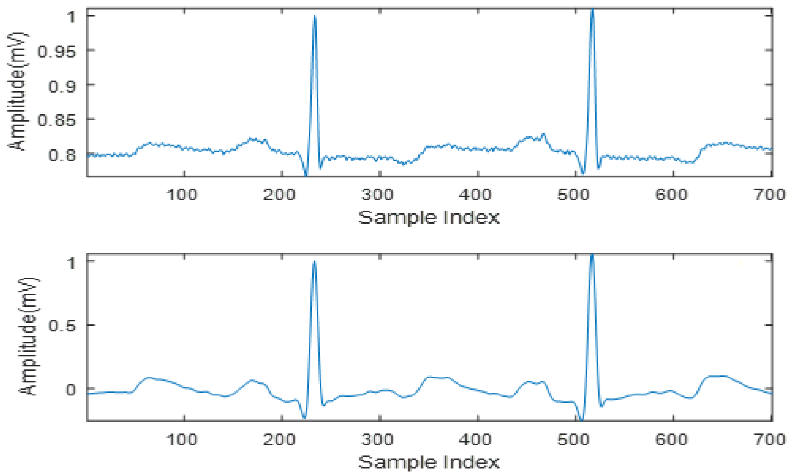


**Fig. 2** *ECG beat before and after preprocessing.*

After the ECG signals were filtered, signals are ready for the segmentation process. The methodology in this paper used the ECG beat segmentation method proposed by Alqudah [8]. The first step of segmentation is the ECG signal normalization, followed by calculating the cumulative area of the ECG signal. This method has successfully extracted 9,100 beats (975 APB, 1300 LBB, 1950 Normal, 3575 PVC, and 1300 RBB). Fig. 3 shows an example of a segmented ECG beat.

## 2.3 Bispectrum

Bispectrum is one of the higher-order spectral analysis of signals. It quantifies the degree of quadratic phase coupling (QPC) and nonlinearity interactions in non-stationary signals. Several types of medical signals are non-stationary signals, such as ECG, EEG, EMG, etc. The first moment of the signal $\mathbf{x}(t)$ is called mean and

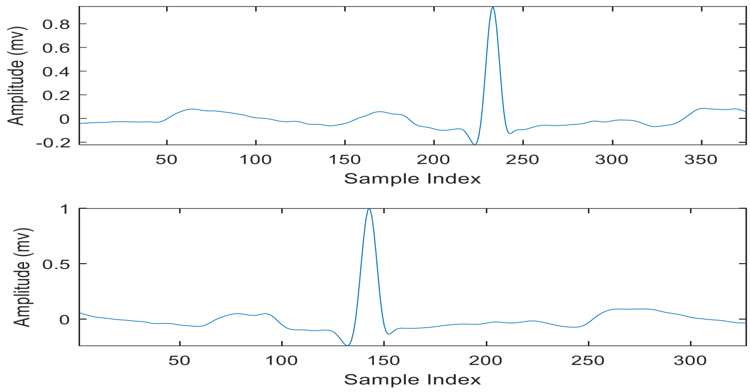


**Fig. 3** *Examples of segmented ECG beats.*

is described by the estimation equation [10]:

$$\mu = E(\mathbf{x}(t))$$

The second moment of the signal is characterized by the equation:

$$\sigma^2 = E(\mathbf{x}(\mathbf{x}(t)(\mathbf{x}(t + t_1))))$$

The third cumulants of the signal are formalized with the equation:

$$\gamma(t_1, t_2) = E(\mathbf{x}(t(\mathbf{x}(t + t_1)\,\mathbf{x}(t_1 + t_2))))$$

The second moment is called the variance of the signal under the assumption of mean equals zero and $t_1 = 0$. On the other hand, the third moment represents the skewness of the signal when both $t_1$ and $t_2$ equal zero. Discrete Fourier Transform of the signal second moment is called the power spectrum. This describes the stationary signals. While DFT of the third moment is named by bispectrum. This quantifies the nonlinearity interaction of the non-stationary signals. Bispectrum is defined by equation [11]:

$$\mathbf{B}(f_1, f_2) = |\mathbf{X}(f_1)\,\mathbf{X}(f_2)\,\mathbf{X}^*(f_1, f_2)|,$$

where $\mathbf{B}(f_1, f_2)$ is the bispectrum at $f_1$ and $f_2$. $\mathbf{X}(f_1)$, $\mathbf{X}(f_2)$, and $\mathbf{X}^*(f_1, f_2)$ are the DFT at $f_1$, $f_2$ and $f_1 + f_2$ respectively.

The squared magnitude of the normalized bispectrum is called the bicoherence. It quantifies the extent phase coupling of the signal and is expressed by the equation [10, 11]:

$$\mathbf{B}_{\mathrm{norm}}(f_1, f_2) = \frac{E(\mathbf{X}(f_1)\,.\mathbf{X}(f_2)\,.\mathbf{X}^*(f_1, f_2))}{\sqrt{\mathbf{P}}(f_1)\,.\mathbf{P}(f_2)\,.\mathbf{P}^*(f_1, f_2)}$$

The bispectrum and the third cumulates are computed for all beats, and the resultant images are stored as image databases with their corresponding categories where the deep learning algorithms are applied on the stored images as it is explained in the next sections in details. Fig. 4 shows the bispectrum general pattern for each of the ECG arrhythmia classes understudy, while Fig. 5 shows the cumulants of the same signals.

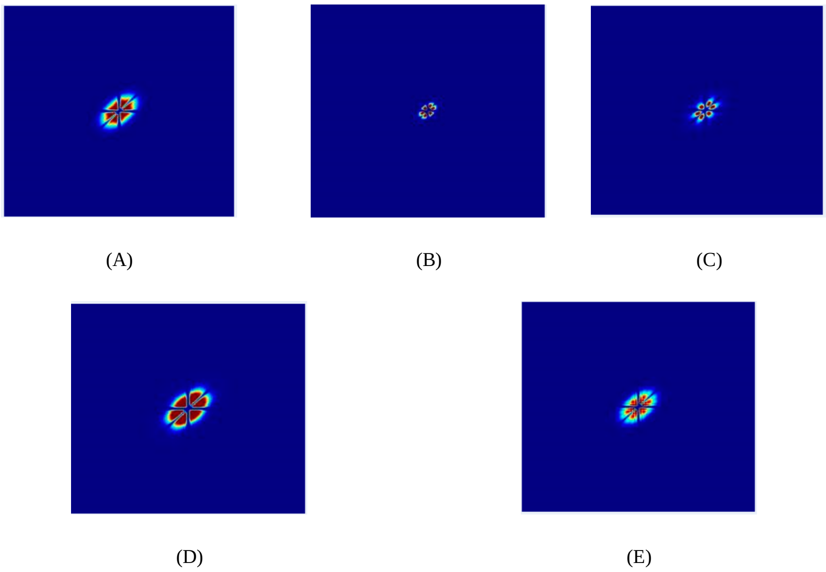


**Fig. 4** *(A) Bispectrum of normal class. (B) Bispectrum of LBB class. (C) Bispectrum of RBB class. (D) Bispectrum of APB class. (E) Bispectrum of PVC class.*

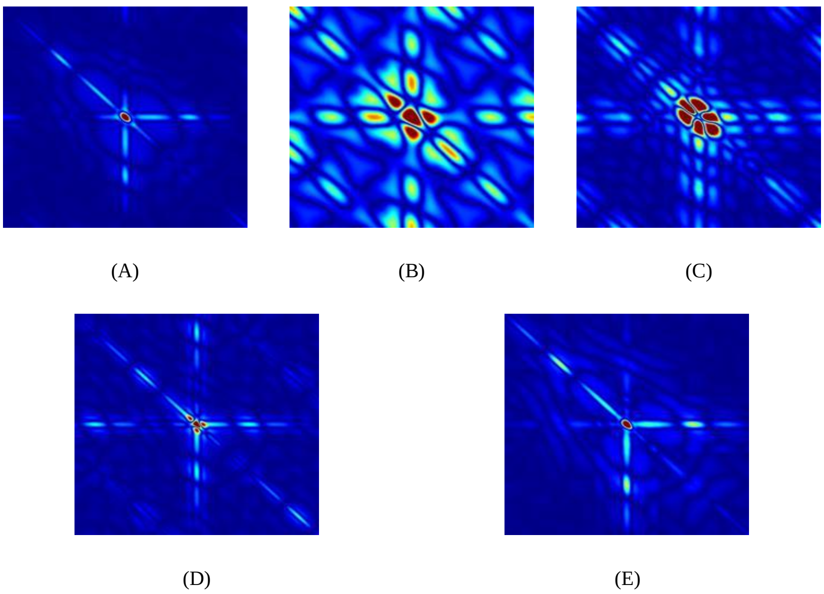


**Fig. 5** *(A) Third cumulants of normal class. (B) Third cumulants LBB class. (C) Third cumulants of RBB class. (D) Third cumulants of APB class. (E) Third cumulants of PVC class.*

## 2.4 Convolutional Neural Network

Convolutional neural network (CNN) is a deep, feed-forward artificial neural network which deals with several types of inputs like 2D images or 1D signals. It consists of an input layer, convolution layer, RELU layer, fully connected layer, classification layer, and output layer [13]. CNN is based mainly on two processes; the first one is convolution, and the second one is down sampling. The convolution processes are performed using trainable filters with pre-specified size and the model weights are adjusting during training phase [14]. In our study, the bispectrum and the third cumulants are computed to all beats, and the resultant images are stored as a database. Five folders are created, each one consists of the generated images for specific class N, LBB, RBB, PVC, and APB. The database is partitioned into training and testing data, 70 % of the data is utilized in the training stage and the rest is used in the test stage. The pre-trained networks are employed in this paper, two pre-trained CNN's that have been used in this paper, namely AlexNet and GoogleNet. The next sections explain the structure of each network separately.

i. AlexNet
   It is one of most used pre-trained convolutional neural networks where it has been pre-trained on more than one million images from the ImageNet database. It consists of 25 layers, starting from the input layer with image size ($227 \times 227 \times 3$), the second one is the convolutional layer with 96 filters and striding [13]. The number of convolutional layers is 5, with a different number of filters and strides. Each Conv layer is followed by RELU layer which is based on $\max(0, x)$ function and it has linear and nonlinear part with 7 layers. The CNN is supported by max pooling, drop out, SoftMax, fully connected layer and ended by classification layer [14].

ii. GoogleNet
   It is a pre-trained convolutional neural network, the main block in the GoogleNet is the inception module. It consists of a parallel of filters with different sizes, that is started from $1 \times 1$ filter to reduce the dimension and to reduce the variance in the image to $3 \times 3$ and $5 \times 5$ and so on. The output from all filters is contaminated [13]. The main distinguishing feature of GoogleNet' is using network-in-network (NIN) approach to enhance and to increase the learning capacity of the convolutional layers besides to dimensionality reduction. NIN structure consists of the main network is sub-networks, where each one is a complete convolutional network with various scales to capture complex behavior. It uses a $1 \times 1$ convolutional kernel to reduce the number of features. The NIN is called inception module. Fig. 6 describes the main component of inception module. Each element of the inception module applies filters which extract features at various scales [16].

The input from the previous layer to the inception layer is distributed to all branches for applying convolution filters with various scales. The filter $1 \times 1$ is used to downsample the numbers of features map, then applying $3 \times 3$ and $5 \times 5$ convolutions with $3 \times 3$ pooling layer with two types of layers, namely, average and max layers [13, 16].

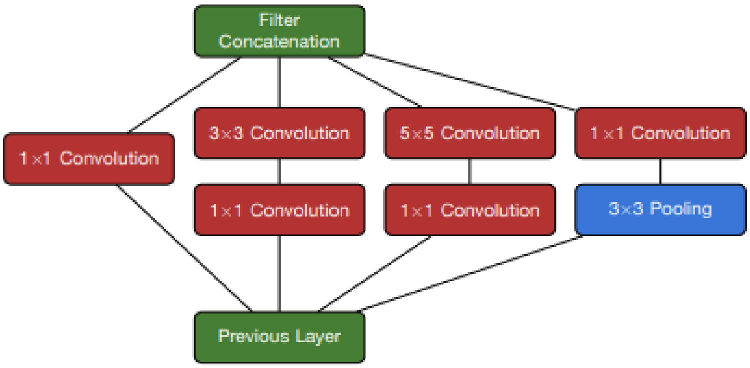


**Fig. 6** *The Diagram of the Inception module [14].*

### 2.5 Transfer learning

Training the entire convolutional neural network by medical images to achieve a certain classification is impossible sometimes due to lack of large medical image dataset. Therefore, random initialization of weights is replaced using exist ConvNet that has been trained on a large dataset [13]. This procedure is called transfer learning, it has two ways; the first is fixed feature extraction and uses the output from an intermediate layer to train the whole classifier. The second one is fine-tuning way, which is based on replacing the last fully connected layer to the new classes and retrain the whole CNN using data that are based on the new modification. This modifying will update the weights in the deeper layers [17].

In this paper, the pre-trained net is modified to be compatible with the goal of this research. The last four layers (pool5-drop_7x7_s1, loss3-classifier, prob, output), are removed and new layers are added to address the number of classes. These layers are (dropout Layer, fully Connected Layer with a number of classes equal to five, and classification Layer) [13, 17].

The optimization method that has been used in this paper is one of the stochastic gradient descent methods, which is based on updating the learning rate. It is called Adaptive moment learning rate (Adam), where it adapts learning rate for each parameter with momentum [13]. The hyperparameter values that have been specified before training both types of classifiers are 0.0001, 64, 20, 0.9000, 0.9900, $1\times10^{-8}$, $1\times, 10F^{-4}$ and infinity for the parameters of Initial learning rate, the mini-batch size, maximum number of epochs, Gradient Decay Factor, Squared Gradient Decay Factor, Epsilon, L2Regularization, and the Gradient Threshold respectively.

## 3. Result

In this research, five types of ECG arrhythmias are classified, namely, Normal beat (N), Right bundle branch block beat (RBBB), Left bundle branch block beat (LBBB), Premature ventricular contraction (PVC), and Atrial Premature beat (APB). The number of the studied ECG beats, which are obtained from the MIT-BIH Arrhythmia database, is 9100 training and testing beats, and they are as follows: 1950 Normal beats, 1300 LBBB beats, 975 APB beats, 3575 PVC beats, and 1300 RBB beats.

After applying the bispectrum and the third cumulants algorithms on each beat individually and saving the resultant images in a database, the resultant images are classified using a pre-trained convolutional neural network. The performance of using AlexNet cumulants are shown in Fig. 7 as a confusion matrix and as training and validation behavior with the iteration number. The overall accuracy of using the third cumulants is 97.6 % for validation and 98.1 % for training. The average accuracy is 97.1 % with sensitivity 94 %, specificity 99.3 %, precision 97.2 %, and false-positive rate 0.7 %, and the overall error is 2.9 %.

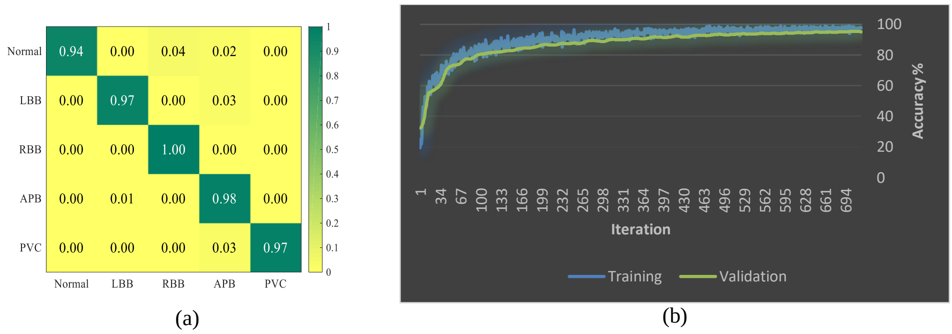


**Fig. 7** *(a) and (b) are confusion matrix and the performance of the training and validation of AlexNet for third cumulants images, respectively.*

By employing bispectrum with AlexNet CNN, the overall accuracy for training was 92.9 % and 91.3 % for validation. Fig. 8 shows the performance of the confusion matrix and the training and validation behavior with the iteration number. The average accuracy is 92.9 %, with sensitivity 92.9 %, specificity 98.2 %, precision 92.9 %, and false-positive rate 1.7 %, and the overall error is 7 %.

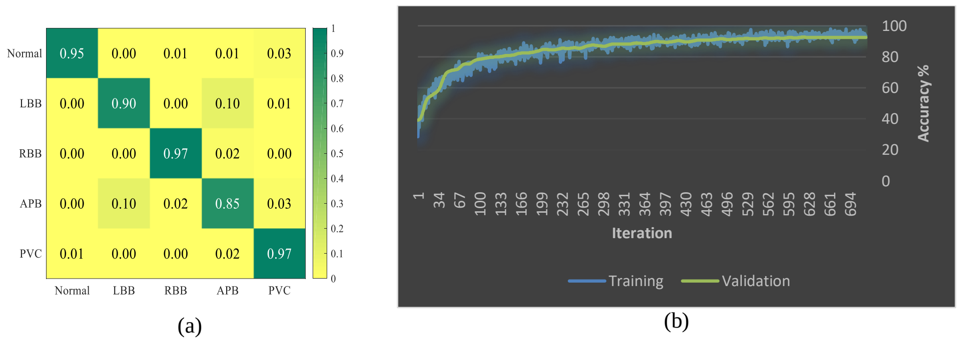


**Fig. 8** *(a) and (b) are confusion matrix and the performance of the training and validation of AlexNet for bispectrum images respectively.*

The training accuracy of applying cumulants with GoogleNet convolutional neural network is 98.2 %. Fig. 9 describes the training and validation process for 10 epochs with batch size 50. The average accuracy is 97.8 %, with sensitivity 97.8 %, specificity 99.4 %, precision 97.8 %, false positive rate 0.57 %, and the overall error is 2 %.

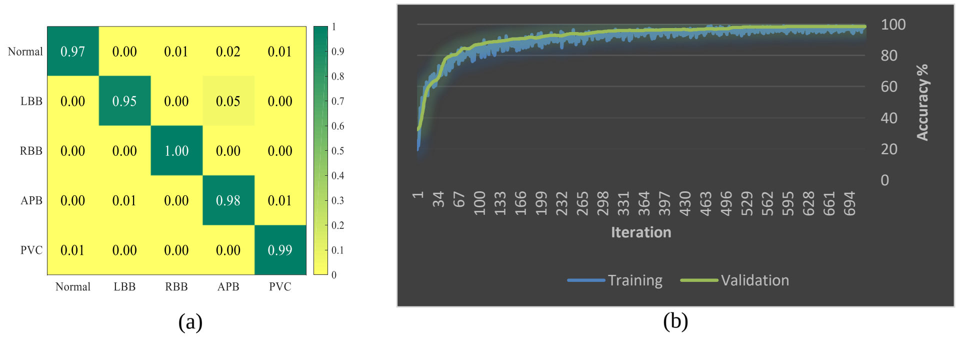


**Fig. 9** *(a) and (b) are confusion matrix and the performance of the training and validation of GoogleNet for Third cumulants images, respectively.*

The bispectrum resultants images are classified into their corresponding class using pre-trained GoogleNet CNN. Fig. 10 a represents the confusion matrix with the overall accuracy for training and validating as 96 % and 95 %, respectively. Fig. 10 b illustrates the behavior of the classifier during the training and testing stage as a function of epochs number. The classifier reaches maximum accuracy at epoch 10. The average accuracy is 94 %, with a sensitivity 94 %, specificity 98.5 %, precision 94.6 %, the false-positive rate of 1.5 %, and the overall error is 2.2 %.

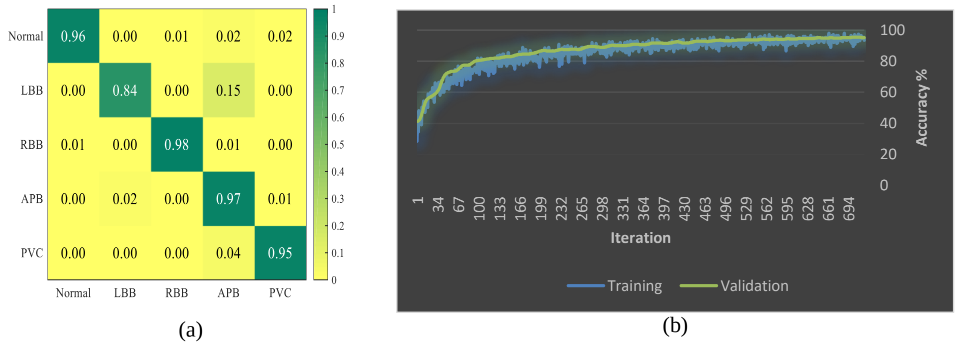


**Fig. 10** *(a) and (b) are confusion matrix and the performance of the training and validation of GoogleNet for bispectrum images, respectively.*

# 4. Discussion and conclusion

In this paper, we proposed an effective five ECG arrhythmias classification method using two-dimensional pre-trained convolutional neural networks with ECG bispectrum and third-order cumulants images as an input. The output colored images with size $361 \times 361 \times 3$ are resized to $224 \times 224 \times 3$ to be compatible with GoogleNet and resized to $227 \times 227 \times 3$ to be compatible with AlexNet as well. These images are generated from the MIT-BIH arrhythmia ECG recording database, where 9100 ECG beat images are obtained with five different classes of ECG beats including normal beat and four abnormal arrhythmia beats.

In Fig. 11 we have compared the results between two types of networks with the specified inputs. It results demonstrate that using the third-order cumulant is better than using bispectrum in specificity, sensitivity, and precision when it is utilized with the AlexNet with a statistical performance of 99.2 % for specificity, 97.1 % for sensitivity and 97.2 % for precision. On the other hand, using bispectrum gives lower specificity as 98.2 %, and sensitivity and precision as 92.2 %. The results are enhanced using GoogleNet Classifier but with better results when using the third-order cumulants algorithm as compared when using the bispectrum algorithm.

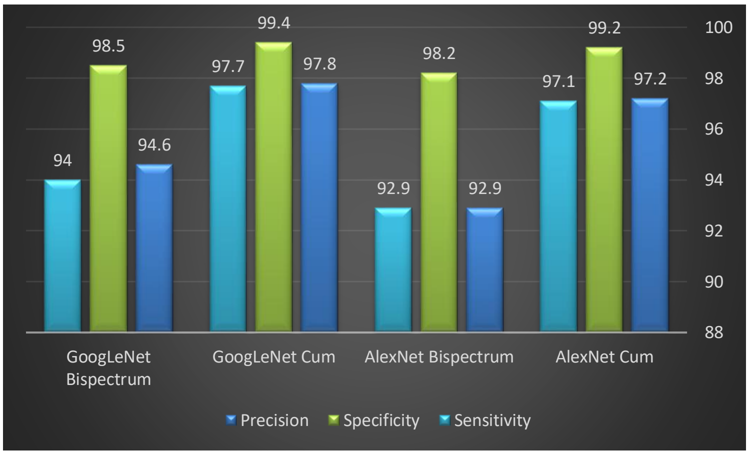


**Fig. 11** *Comparison between two CNN's with third cumulants and bipectrum.*

As aforementioned, the GoogleNet classifier results are better than using AlexNet classifier, and this is due to the architecture of GoogleNet, which has an inception module structure. This structure enables the network to extract features more accurately than using AlexNet structure, which is based on series convolutional layers.

Because of the nature of ECG signal which can be modeled as a gaussian functions [8], the results were obtained using the third-order cumulant is better than employing the bispectrum, where the power spectrum in bispectrum algorithm suppresses all phase relations in the signal and this type of algorithms is used mainly for detecting and quantifying the phase coupling as in analyzing a non-Gaussian signal found in EEG [11]. Therefore, the classification results are better in case of using the third-order cumulants algorithm. Our ECG arrhythmia classification result indicates that detection of arrhythmia with ECG images and CNN model can be an effective approach to help the experts to diagnose cardiovascular diseases which can be monitored using ECG signals. Furthermore, the proposed ECG arrhythmia classification method can be exploited in many fields such as in medical robots or in ECG signal monitors to help cardiologists in identifying arrhythmias more precisely and efficiently in a real-time manner.

**Conflict of interest** The authors declare no conflict of interest.

## Acknowledgement

The authors would like to acknowledge the authors of the dataset from PhysioNet used in this research paper and the anonymous reviewers for valuable comments.

# References

[1] ACHARYA U.R., at al. A deep convolutional neural network model to classify heartbeats. *Computers in biology and medicine*. 2017, 89, pp. 389–396, doi: 10.1016/j.compbiomed.2017.08.022.

[2] BANERJEE S., MITRA M. Application of cross wavelet transform for ECG pattern analysis and classification. *IEEE transactions on instrumentation and measurement*. 2013, 63.2, pp. 326–333. doi: 10.1109/TIM.2013.2279001.

[3] JAMBUKIA S., DABHI V., PRAJAPATI H. Classification of ECG signals using machine learning techniques: A survey. In: *2015 International Conference on Advances in Computer Engineering and Applications*. IEEE, 2015, pp. 714–721. doi: 10.1109/ICACEA.2015.7164783.

[4] SHEN C., at al. Detection of cardiac arrhythmia in electrocardiograms using adaptive feature extraction and modified support vector machines. *Expert Systems with Applications*, 2012, 39.9, pp. 7845–7852. doi: 10.1016/j.eswa.2012.01.093.

[5] KACHUEE M., FAZELI S., SARRAFZADEH M. Ecg heartbeat classification: A deep transferable representation. In: *2018 IEEE International Conference on Healthcare Informatics (ICHI). IEEE*, 2018, pp. 443–444. doi: 10.1109/ICHI.2018.00092.

[6] JOSHI N., TOPANNAVAR P. Support vector machine based heartbeat classification. In: Proc. of *4th IRF Int. Conf.2014*, pp. 140-144. ISSN: 2321–9009

[7] FIROOZABADI R., GREGG R., BABAEIZADEH S. Identification of exercise-induced ischemia using QRS slopes. *Journal of electrocardiology*, 2016, 49.1, pp. 55–59. doi: 10.1016/j.jelectrocard.2015.09.001.

[8] ALQUDAH A. An enhanced method for real-time modelling of cardiac related biosignals using Gaussian mixtures. *Journal of medical engineering & technology*, 2017, 41.8, pp. 600–611. doi: 10.1080/03091902.2017.1382587.

[9] GOLDBERGER Ary L. at al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 2000, 101.23, pp. e215–e220. doi: 10.1161/01.CIR.101.23.e215.

[10] GOSHVARPOUR A., at al. Bispectrum estimation of electroencephalogram signals during meditation. *Iranian journal of psychiatry and behavioral sciences*, 2012, 6.2, pp. 48–54. PubMed PMID: 24644482; PubMed Central PMCID: PMC3940018.

[11] SANDERS P. at al. Spectral analysis identifies sites of high-frequency activity maintaining atrial fibrillation in humans. *Circulation*, 2005, 112.6, pp. 789–797. doi: 10.1161/CIRCULATIONAHA.104.517011.

[12] ALQUDAH A. at al. Developing of robust and high accurate ECG beat classification by combining Gaussian mixtures and wavelets features. *Australasian physical & engineering sciences in medicine*, 2019, 42.1, pp. 149–157. doi: 10.1007/s13246-019-00722-z.

[13] ALQUDAH A., ALQURAAN H., QASMIEH I. Segmented and Non-Segmented Skin Lesions Classification Using Transfer Learning and Adaptive Moment Learning Rate Technique Using Pretrained Convolutional Neural Network. In: *Journal of Biomimetics, Biomaterials and Biomedical Engineering*. Trans Tech Publications Ltd, 2019, pp. 67–78. doi: 10.4028/www.scientific.net/JBBBE.42.67.

[14] LIU T. at al. Implementation of training convolutional neural networks. arXiv preprint arXiv:1506.01195, 2015.

[15] ALOM M., at al. The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint* arXiv:1803.01164, 2018.

[16] SZEGEDY C, at al. Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1–9. doi: 10.1109/CVPR.2015.7298594.

[17] HUANG Z., PAN Z., LEI B. Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. *Remote Sensing*, 2017, 9(9). doi: 10.3390/rs9090907.