A Novel Technique in Classifying Heart Diseases based on Electrocardiogram (ECG) Signals using Deep Learning and Spectrogram Image Analysis

Edward B. Panganiban[1], Arnold C. Paglinawan[2], Wen Yaw Chung[3], Gilbert Lance S. Paa[4]

[1]College of Computing Studies, Information and Communication Technology, Isabela State University
Echague, Isabela, Philippines, ebpanganiban@isu.edu.ph
[2]School of Electrical, Electronics and Computer Engineering, Mapua University, Manila, Philippines,
acpaglinawan@mapua.edu.ph
[3]Department of Electronic Engineering, Chung Yuan Christian University, Chung-Li, Taiwan, R.O.C.,
eldanny@cycu.edu.tw
[4]Fellow Philippine College of Physician, member – Philippine Heart Association and Philippine Society of
Vascular Medicine, gilbertlancepaa@gmail.com

ABSTRACT

Deep learning is among the most reliable approaches in analyzing physiological signals for healthcare applications. This makes data mining strategies more useful and reliable in analyzing a large group of data like medical records. Through this technology breakthrough, the authors established a novel approach in examining electrocardiogram (ECG) records and classify them on what kind of heart disease based from its pattern using spectrogram image analysis. These records underwent data training as part of the deep learning process. Data trained are in the form of extracted ECG logs in comma-separated values (CSV) format from different patient cases. These databases are downloaded from Physionet databases, which offers free access to an extensive collection of recorded physiological data. Once done downloading, these databases were categorized per disease type. Records of diseases that were in the same category must be saved in the same folder. These CSV files passed through a preprocessing procedure through filtering CSV log files and cut each ECG reading into 10-second duration and save it as text file. These text files were converted into their spectrogram images equivalent. The converted text files to spectrogram images were still categorized per disease type. The file name of each folder containing the spectrogram images for each disease type served as the label map for the output classifier. The final layer was retrained for dataset development through the use of Google's Inception V3 model that serves as a base learning model. Fine tuning was established through setting its training rate to 0.001, training batch size to 500, a validation batch size of 200 and a total number of training steps of 500,000. These settings had achieved the optimum result of 100% validation and training accuracy each. The result of final test accuracy is 98.9% and had generated a frozen model for the classifier. Its accuracy can improve if there are more available and reliable ECG records. Using the frozen model, the same procedures were done in preparing the dataset to be used for the testing stage. The result of the classification of disease is shown through the percentage value of how likely is the input spectrogram image based from disease categories. The authors evaluated the proposed technique and obtained a significant performance on the classification of types of heart diseases based on available databases from Physionet.

Keywords: deep learning, ECG, spectrogram, classification, heart disease, Physionet, inception v3

1. INTRODUCTION

According to National Vital Statistics Reports, the leading death reason in 2016 is cardiovascular or heart diseases. About 23% of the total deaths is caused by heart disease [1]. In this paper, the available heart disease databases from the physionet databank were collected for experimentation. Physionet, a web-based resources of physiological signals, is a free access for signals archiving and signal processing [2]. The authors examined and collected several heart diseases' databases but consider four of these initially because of its reliability and electrocardiogram (ECG) property. These are atrial fibrillation (AF), bradycardia, tachycardia, and bundle branch block. Healthy signals were also collected and included in the experiment since these are very important physiological signals. These physiological datasets are in the form of ECG signals. Electrocardiograms (ECG), discovered by Muirhead in 1872 via wires attached to the patient's wrist, are used to obtain a record of a patient's heartbeat. This is a widely used non-invasive approach for clinical diagnosis in patients with a cardiovascular disease [3]. Checking heart's electrical activity needs to have in-depth skills and professional knowledge to accurately explain ECG [4]. Furthermore, visually inspect ECG signals takes more time compared to today's technological approaches [5].

Atrial fibrillation is an irregular heartbeat called arrhythmia which may lead can lead to various heart-related conditions such as blood clots, stroke, and heart failure [6]. Bradycardia is a condition of a patient that has a very slow heart rate of less than 60 beats per minute while tachycardia has a very fast heart rate with at least 100 beats per minute [7]. A bundle branch block is a disease wherein some part of the heart's conduction system is "blocked", specifically, an impulse is blocked as it travels through the bundle branches [8]. These diseases can be determined through the characteristics of ECG signals.

In this paper, the authors established a novel approach in determining the type of heart disease without the need for ECG visual inspection like P-peak or R-peak detection [5]. This is through a spectrogram process and deep learning analysis using the csv file form of ECG signals to classify the type of heart disease. This paper utilizes the Inception V3 model in TensorFlow platform for the transfer learning technology to retrain the heart diseases' datasets. This is an open source algorithm made by Google to classify images [9]. TensorFlow [10], established by Google, is a second generation artificial intelligence learning system. It is compatible with neural network models like a recurrent neural network (RNN) and convolutional neural network (CNN). This model is utilized for image and speech recognition and other machines deep learning applications. This study has established and developed an image classification model with better efficiency and higher accuracy.

A convolutional neural network (CNN) [9],[11] has been utilized recently due to its efficient recognition capability and great performance in the image classification field. CNN played an important role in data mining and machine learning techniques, specifically Deep Learning Networks (DLNs). These DLNs are recent data analysis tools used for healthcare diagnosis based on physiological data according to related articles [12]–[18]. Furthermore, previous studies showed that DLNs can help physician in analyzing medical images with higher accuracy than radiologists [19]. This powerful process of DLNs is primarily given the fact that they can identify characteristic features of a given datasets during its learning process to make accurate predictions. With numerous novel approaches in medical applications using DLNs, the excessive computational cost from the conventional processes was addressed.

Deep learning [20] is among the most reliable approaches in analyzing physiological signals for healthcare applications. This makes data mining strategies more useful and reliable in analyzing a large group of data like medical records. Through this technology breakthrough, the authors established a novel approach in examining electrocardiogram (ECG) records and classify them on what kind of heart disease based from its pattern using spectrogram image analysis. These records underwent data training as part of the deep learning process.

2. METHODOLOGY

2.1 Conceptual framework

The novel approach in classifying heart diseases based on ECG is illustrated through the conceptual framework given in figure 1. The framework has three major stages that comprise of preprocessing, the training process and the testing process. These stages have subroutines that prepare and analyze input data to generate the correct output.

2.1.1. Preprocessing

In this stage, it requires extracted ECG logs in csv form from different patient cases. The databases used in this paper are downloaded from https://physionet.org/ wherein datasets were identified and serves as a model database. Once done, the databases must be categorized per disease type. Diseases that are in the same category must be saved in the same folder. The csv files are subjected preprocessing procedure through filtering the csv log files and cut each ECG reading in 10 second duration and save it as a text file. These text files are converted into its equivalent spectrogram images through python codes. Spectrogram images should remain categorized per disease type. The file name of each folder containing the spectrogram images in each type of disease serves as the label map for the output classifier. Thus, it requires to be precise.

2.1.2 Training Process

Using Google's Inception v3 as a base learning model, the final layer is retrained to create its own dataset. To attain its fine tuning, the training rate setting must be 0.01, the number of training steps can be up to 500,000, training batch size set to 100, and should have a validation batch size of 100. Once the training is completed and has achieved the optimum accuracy, a frozen model is then generated for the classifier and is ready to deploy. This frozen model will be used for the classification of test datasets.

2.1.3 Testing Process

The last stage is for the testing and validation of the frozen model output. It performs a classification process wherein it requires the same procedure with the training stage when preparing for the test datasets. In testing the accuracy of the classifier, the test datasets were taken from the physionet website but these ECG logs were not part of the training datasets. These csv files were also subjected to filtering/staging process through dicing the csv file into text files with 5 minutes duration of ECG reading. The text file is converted to spectrogram images and was fed in the frozen model for the classification of disease. The classification output is displayed in descending order based on the percentage of how likely is the input spectrogram image as compared to the categorized database.
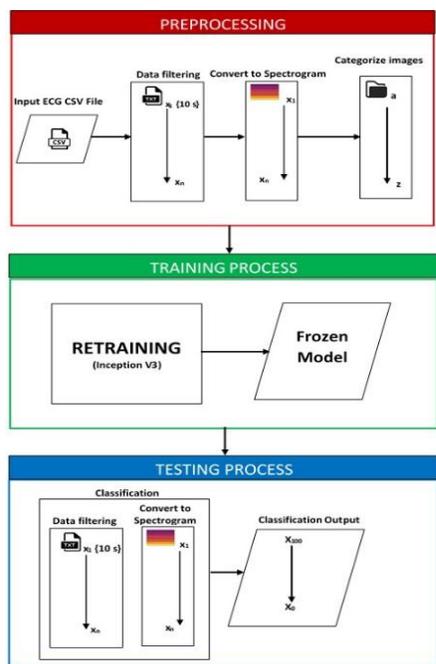
Figure 1. Conceptual Framework

Figure 2. ECG signals (a. healthy b. tachycardia c. bradycardia d. bundle branch block e. atrial fibrillation)

3.2 Spectrogram process

This process requires ECG input that is in txt format. The spectrogram process is done through lines of python codes to produce images such as what are shown in figure 3. The figures are 25 example output images out of 1048 generated spectrogram images of healthy signals.
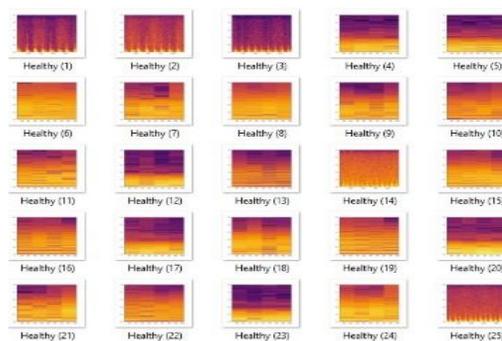

Figure 3. Sample spectrogram output images (healthy)

Figure 4 are samples of images taken from the spectrogram process performed in tachycardia (figure 4a) and bundle branch block (figure 4b). There is a total of 1067 images of tachycardia and 82 images of bundle branch block. In here, patterns are already noticeable.
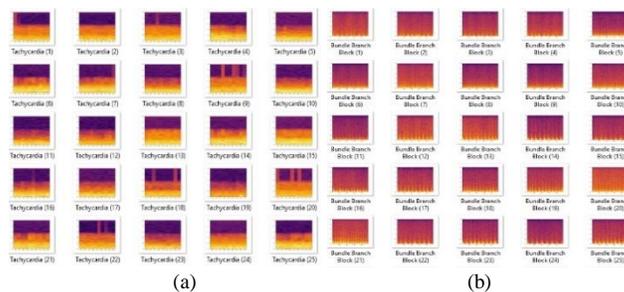
2.2 Training datasets

Table 1 is the list of databases used in this paper indicating the number of csv files downloaded from the physionet databank. The databases related to heart diseases includes atrial fibrillation, bradycardia, bundle branch block and tachycardia. Based from the table, the spectrogram images produced from atrial fibrillation datasets resulted to 265 image data, 1426 for bradycardia, 82 for bundle branch block and 1067 for tachycardia. Healthy csv logs of 121 produced a total of 1408 spectrogram images.

Table 1: Training datasets

Database	Number of datasets in CSV form	Spectrogram images
Atrial Fibrillation	105	265
Bradycardia	47	1,426
Bundle Branch Block	15	82
Tachycardia	35	1,067
Healthy	121	1,408

3. RESULTS

3.1 Data Filtering

Input data are in the form of csv file as shown in figure 2. These are sample ECG signals of healthy, tachycardia, bundle branch block, bradycardia and atrial fibrillation taken from physionet with 10 s duration. These are converted into txt format in preparation for the spectrogram process.

(a)   (b)

Figure 4. Sample spectrogram output images (a. tachycardia b. bundle branch block)

Finally, figure 5 is the produced images of bradycardia and atrial fibrillation executed by the spectrogram process. There are 1,426 images generated for bradycardia and 265 images for atrial fibrillation. Pattern analysis can already be performed from these images.
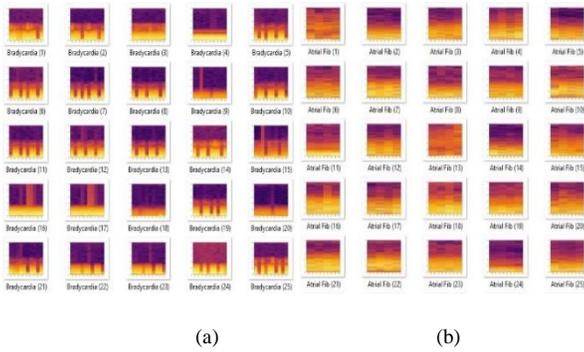
(a)   (b)

Figure 5. Sample spectrogram output images (a. bradycardia b. atrial fibrillation)

3.3 Deep Learning

After spectrogram images are generated, these are categorized according to its type of database. In this paper, the authors used these heart diseases and healthy databases as an input for the deep learning model. The learning environment used is the Inception v3 algorithm of the tensor flow platform. This transfer learning method has a powerful capability to retrain datasets with very high classification accuracy. The Inception v3 model main graph is illustrated in figure 6. This is a deep neural network model that requires a GPU-based computer for the training process. It takes at least few days to train datasets. For a 500,000 steps of training, it needs at least 3 days to finish. The final layer of the model is trained using back propagation algorithm and its cost entropy function is utilized to regulate the weight parameter through computing the error between the output of the softmax layer and the database type. For its optimum performance, the researchers found out that the setting for its fine tuning must have a training rate of 0.001, a training batch size of 500, and a validation batch size of 200.
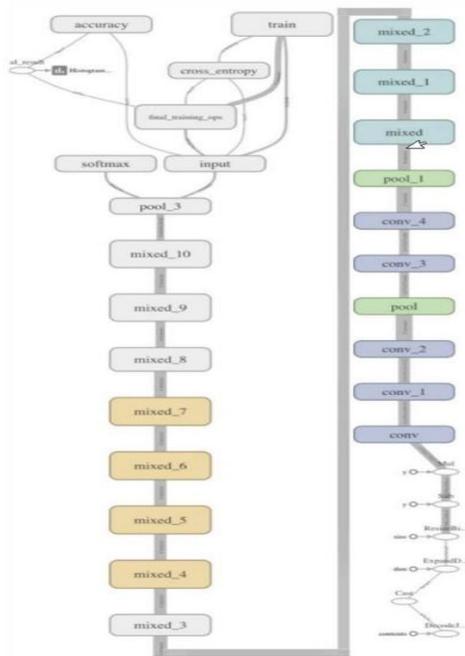

Figure 6. Main graph of Inception v3 model

3.4 Experimental Results

This paper performed four experiments in training the datasets. The first experiment was the combination of healthy and tachycardia datasets. Figure 7 shows the results of cross entropy after the retraining the datasets into 4,000 training steps. The first figure located at the top is the complete graph of the train logs while the bottom figure is the zoomed graph near the end of the training steps. It is given in here that the cross entropy cost function or the error level is less than 0.0050 or 0.50%. This value indicates very high classification accuracy. In figure 8, it shows the result of the validation (orange line) and the training (blue line) set. Top figure is the complete logs while bottom figure is the zoomed version. This means that it obtained 100% accuracy for both validation and training.
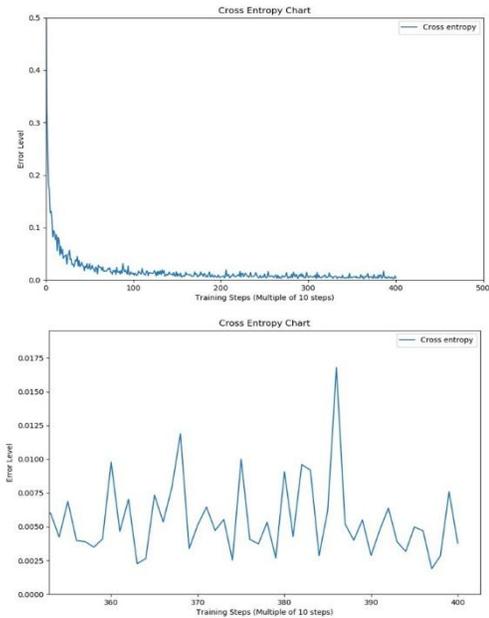

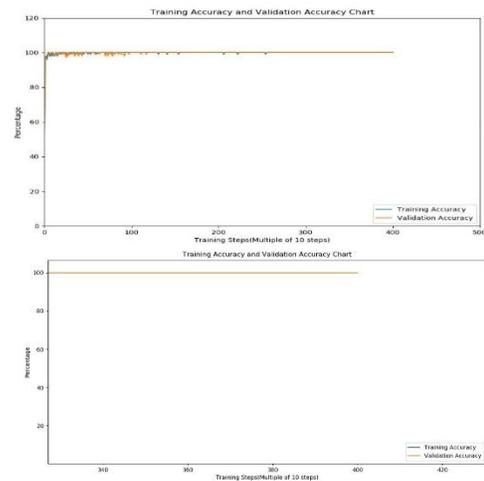**Figure 7.** Cross entropy result (healthy and tachycardia)

Figure 8. Training and Validation results (healthy and tachycardia)

The second experiment is the combination of healthy, tachycardia and bundle branch block which underwent 30,000 training steps. Figure 9 shows the result of the cross entropy while figure 10 is for the training and validation accuracy. The final cross entropy cost is only 2.16% with a validation and training accuracy of 100%.
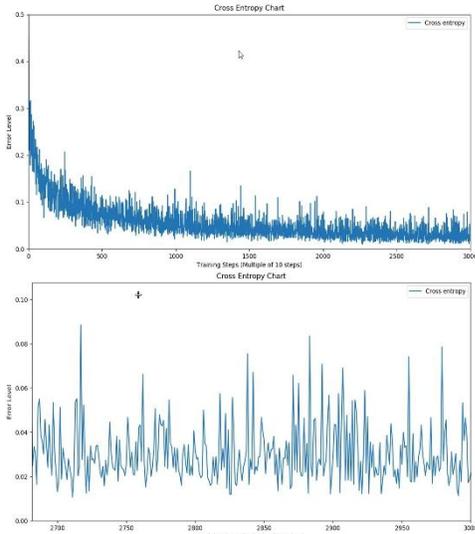

Figure 9. Cross entropy result (Healthy, Tachycardia and Bundle Branch Block)
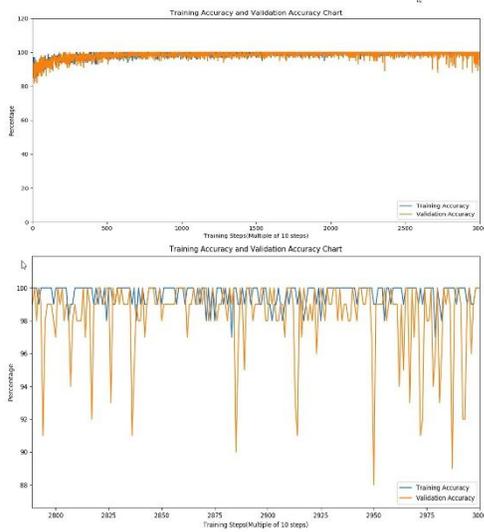

Figure 10. Training and Validation results (Healthy, Tachycardia and Bundle Branch Block)

Third experiment is the combination of healthy, tachycardia, bundle branch block and bradycardia. The result of the cross entropy shown in figure 11 attained is 6.77% only at 50,000 training steps. Validation accuracy arrived at 98% while training accuracy obtained 99% which were shown in figure 12.

Finally, the last experiment is a combination of 5 databases comprising of healthy, tachycardia, bundle branch block, bradycardia and atrial fibrillation. It attained 1.47% error level, or cross entropy cost as shown in figure 13. However, it needs at least 500,000 training steps to attain this. Based on figure 14, the validation and training accuracy resulted to

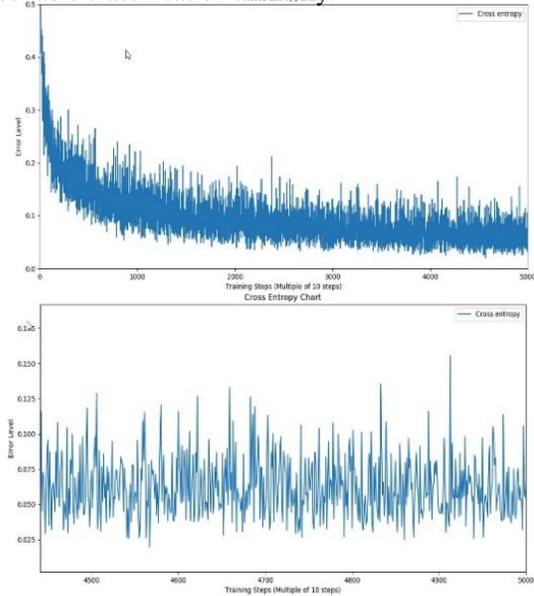100% each. Thus, the frozen model result achieved a desirable classification accuracy.

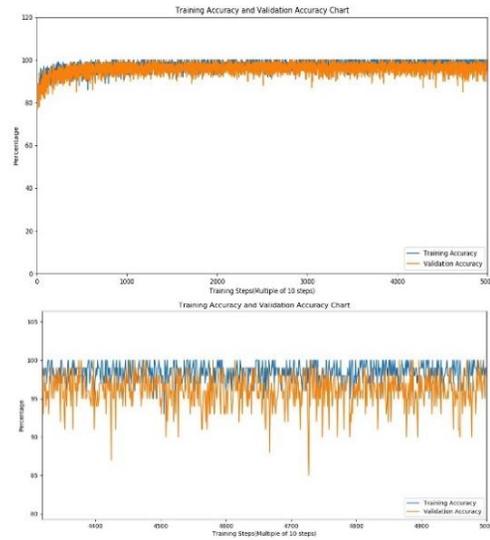**Figure 11.** Cross entropy result (Healthy, Tachycardia, Bundle Branch Block and Bradycardia)

Figure 12. Training and Validation results (Healthy, Tachycardia, Bundle Branch Block and Bradycardia)
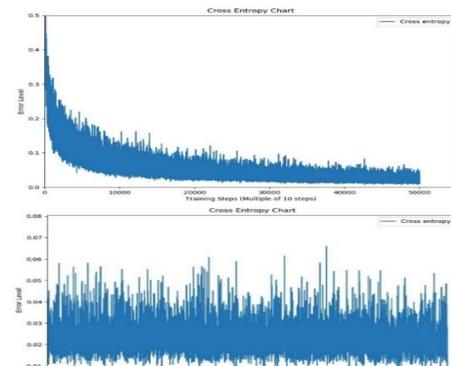

Figure 13. Cross entropy result (Healthy, Tachycardia, Bundle Branch Block, Bradycardia and Atrial Fibrillation)
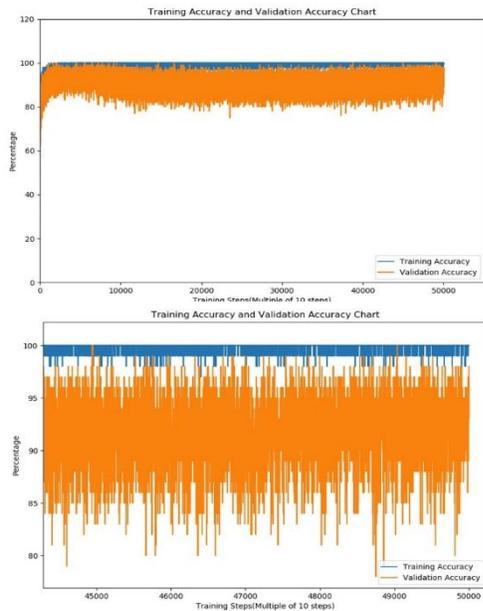
Figure 14. Training and Validation results (Healthy, Tachycardia, Bundle Branch Block, Bradycardia and Atrial Fibrillation)

3.5 Tabulated summary result

Table 2 summarizes the experimental results for every combination of datasets. The final test accuracy decreases when adding datasets, but the error level decreases when the training step is increased. The final test accuracy for the five datasets of ECG signals achieved 98.9% and a minimal error rate of 1.47% only. Therefore, increasing train steps while adding databases will increase its final test accuracy and its cross entropy will be lessened.

Table 2: Summary of results

Database	Train Steps	Final Validation Accuracy	Final Training Accuracy	Final Test Accuracy	Percent Error (Cross entropy)
Healthy and Tachycardia	**4,000**	100%	100%	**100%**	0.32%
Healthy, Tachycardia and Bundle Branch Block	**30,000**	100%	100%	**98.5%**	2.16%
Healthy, Tachycardia, Bundle Branch Block and Bradycardia	**50,000**	98%	99%	**97.9%**	6.77%
Healthy, Tachycardia, Bundle Branch Block, Bradycardia and Atrial Fibrillation	**500,000**	100%	100%	**98.9%**	1.47%

3.6 Testing Result

In order to test the classifier, samples of csv files that do not belong to the training datasets were tested. These were the datasets from those diseases that were separated before the retraining. Figure 15 shows the results of the classifier. All database types were tested to determine the accuracy of the classifier. The researchers found out that the technique is accurate in classifying the specified types of heart disease.
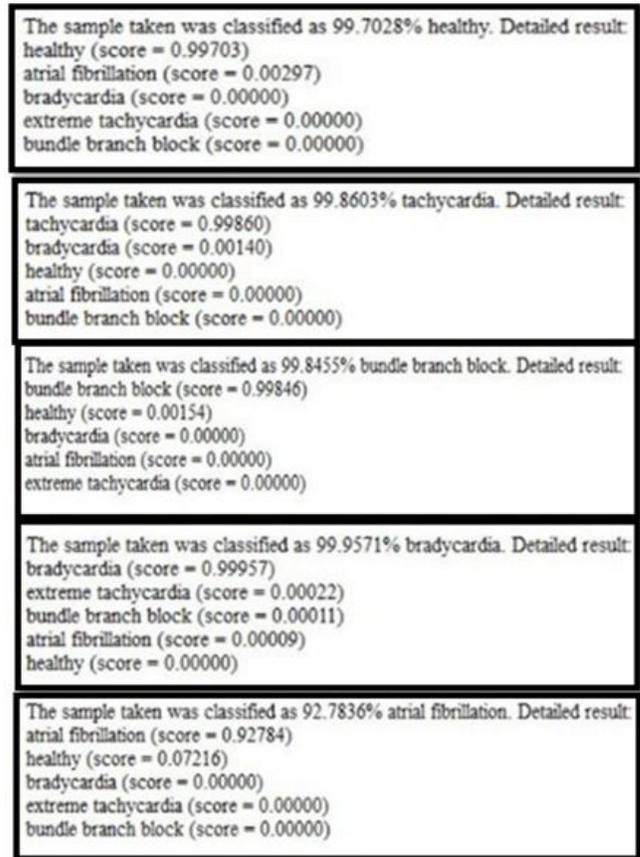

Figure 15. Test results from random samples of datasets

4 CONCLUSION

A new approach in classifying the types of cardiovascular diseases was established using spectrogram process and deep learning in this paper. The datasets, which are ECG signals in csv format from physiobank, served as inputs to this new technique. The datasets under experiments were healthy, tachycardia, bradycardia, bundle branch block and atrial fibrillation. Using spectrogram for preprocessing and Inception v3 for its learning process, its final training and validation accuracy is 100% with a final test accuracy of 98.9%. The cross-entropy loss reached only 1.47% which indicates a very low error rate. Based on the results, adding available and reliable databases of heart diseases in csv form is recommended for further research.

REFERENCES

[1]   B. B. Xu J, Murphy SL, Kochanek KD, "**National Vital Statistics Reports Deaths : Final Data for 2013,**" *Natl Vital Stat Rep*, vol. 65, no. 5, pp. 1–76, 2016.

[2]   G. B. Moody, R. G. Mark, and A. L. Goldberger, "**Physionet: A web-based resource for the study of physiologic signals**," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 70–75, 2001.
https://doi.org/10.1109/51.932728

[3]   Z. Xiong, M. Stiles, and J. Zhao, "**Robust ECG Signal Classification for the Detection of Atrial Fibrillation Using Novel Neural Networks**," *Comput. Cardiol. (2010).*, vol. 44, no. 2017, pp. 1–4, 2017.
https://doi.org/10.22489/CinC.2017.066-138

[4]   J. R. Mehall, R. M. Kohut, E. W. Schneeberger, W. H. Merrill, and R. K. Wolf, "**Absence of Correlation Between Symptoms and Rhythm in 'Symptomatic' Atrial Fibrillation**," *Ann. Thorac. Surg.*, vol. 83, no. 6, pp. 2118–2121, 2007.
https://doi.org/10.1016/j.athoracsur.2007.02.084

[5]   S. Asgari, A. Mehrnia, and M. Moussavi, "**Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine**," *Comput. Biol. Med.*, vol. 60, no. 2015, pp. 132–142, 2015.
https://doi.org/10.1016/j.compbiomed.2015.03.005

[6]   Y. Chang, S. Wu, L. Tseng, H. Chao, and C. Ko, "**AF Detection by Exploiting the Spectral and Temporal Characteristics of ECG Signals with the LSTM Model**," *Natl. Chiao Tung Univ.*, pp. 1–4, 2018.
https://doi.org/10.22489/CinC.2018.266

[7]   D. Cooley, "**Categories of Arrhythmias**," *Texas Hear. Inst.*, pp. 1–8, 2018.

[8]   D. Cooley, "**Bundle Branch Block**," *Texas Hear. Inst.*, vol. 1, no. 1, p. 6, 2019.

[9]   J. Bankar and N. R. Gavai, "**Convolutional Neural Network based Inception v3 Model for Animal Classification**," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 7, no. 5, pp. 142–146, 2018.

[10]   S. S. Girija, "**TENSORFLOW : Large-Scale Machine Learnng on Heterogeneous Distributed Systems**," *Softw. vailable from tensorflow.org*, pp. 1–25, 2016.

[11]   J. R. B. Del Rosario, "**Development of a Face Recognition System using Deep Convolutional Neural Network in a Multi-View Vision Environment**," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 3, pp. 369–374, 2019.
https://doi.org/10.30534/ijatcse/2019/06832019

[12]   K. H. Lee, S. Y. Kung, and N. Verma, "**Low-energy formulations of support vector machine kernel functions for biomedical sensor applications**," *J. Signal Process. Syst.*, vol. 69, no. 3, pp. 339–349, 2012.
https://doi.org/10.1007/s11265-012-0672-8

[13]   W. Karlen, C. Mattiussi, and D. Floreano, "**Sleep and wake classification with ECG and respiratory effort signals**," *IEEE Trans. Biomed. Circuits Syst.*, vol. 3, no. 2, pp. 71–78, 2009.
https://doi.org/10.1109/TBCAS.2008.2008817

[14]   V. Podgorelec, P. Kokol, and I. Rozman, "**Decision Trees: An overview and Tehir Use in Medicine**," *J. Med. Syst.*, vol. 26, no. 5, pp. 445–463, 2002.
https://doi.org/10.1023/A:1016409317640

[15]   T. H. N. Vu, N. Park, Y. K. Lee, Y. Lee, J. Y. Lee, and K. H. Ryu, "**Online discovery of Heart Rate Variability patterns in mobile healthcare services**," *J. Syst. Softw.*, vol. 83, no. 10, pp. 1930–1940, 2010.
https://doi.org/10.1016/j.jss.2010.05.074

[16]   D. Giri *et al.*, "**Automated diagnosis of Coronary Artery Disease affected patients using LDA, PCA, ICA and Discrete Wavelet Transform**," *Knowledge-Based Syst.*, vol. 37, pp. 274–282, 2013.
https://doi.org/10.1016/j.knosys.2012.08.011

[17]   F. Hu, M. Jiang, L. Celentano, and Y. Xiao, "**Robust medical ad hoc sensor networks (MASN) with wavelet-based ECG data mining**," *Ad Hoc Networks*, vol. 6, no. 7, pp. 986–1012, 2008.
https://doi.org/10.1016/j.adhoc.2007.09.002

[18]   J. Y. Yeh, T. H. Wu, and C. W. Tsao, "**Using data mining techniques to predict hospitalization of hemodialysis patients**," *Decis. Support Syst.*, vol. 50, no. 2, pp. 439–448, 2011.
https://doi.org/10.1016/j.dss.2010.11.001

[19]   Abder-Rahman Ali, "**Deep Learning Medical Applications**," *https://www.techemergence.com*, pp. 1–10, 2017.

[20]   R. Biswas and S. Roy, "**Content Based CT image Sign Retrieval using Fast Discrete Curvelet Transform and Deep Learning**," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 3, pp. 854–863, 2019.
https://doi.org/10.30534/ijatcse/2019/80832019