Computers in Biology and Medicine 113 (2019) 103387

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals

Ozal Yildirim[a,*], Muhammed Talo[a], Betul Ay[b], Ulas Baran Baloglu[c], Galip Aydin[b], U. Rajendra Acharya[d,e,f]

[a] Department of Computer Engineering, Munzur University, Tunceli, Turkey
[b] Department of Computer Engineering, Fırat University, Elazığ, Turkey
[c] Department of Computer Science, University of Bristol, Bristol, United Kingdom
[d] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
[e] Department of Biomedical Engineering, School of Science and Technology, Singapore School of Social Sciences, Singapore
[f] School of Medicine, Faculty of Health and Medical Sciences, Taylor's University, 47500, Subang Jaya, Malaysia

ARTICLE INFO

Keywords:
Diabetes mellitus
Heart rate signals
Deep learning
Transfer learning

ABSTRACT

In this study, a deep-transfer learning approach is proposed for the automated diagnosis of diabetes mellitus (DM), using heart rate (HR) signals obtained from electrocardiogram (ECG) data. Recent progress in deep learning has contributed significantly to improvement in the quality of healthcare. In order for deep learning models to perform well, large datasets are required for training. However, a difficulty in the biomedical field is the lack of clinical data with expert annotation. A recent, commonly implemented technique to train deep learning models using small datasets is to transfer the weighting, developed from a large dataset, to the current model. This deep learning transfer strategy is generally employed for two-dimensional signals. Herein, the weighting of models pre-trained using two-dimensional large image data was applied to one-dimensional HR signals. The one-dimensional HR signals were then converted into frequency spectrum images, which were utilized for application to well-known pre-trained models, specifically: AlexNet, VggNet, ResNet, and DenseNet. The DenseNet pre-trained model yielded the highest classification average accuracy of 97.62%, and sensitivity of 100%, to detect DM subjects via HR signal recordings. In the future, we intend to further test this developed model by utilizing additional data along with cloud-based storage to diagnose DM via heart signal analysis.

1. Introduction

Diabetes Mellitus (DM) occurs when the blood glucose level is above normal. Diabetes is a metabolic disorder that can happen at any age and cause serious complications. There are two major types of diabetes: Type 1 and Type 2 [1]. Type 1 DM occurs when there is no insulin in the body or minimal insulin production. Type 1 DM, which is usually seen in children and adolescents (< 30), causes coma, and there is a risk of death if not treated. Type 2 DM is a disease that usually occurs in middle-aged or elderly patients (> 40). It occurs when the produced insulin is not consumed in the body, or sufficient insulin is not produced by the body.

As reported by the International Diabetes Federation, 12% of global health spending is spent on diabetes, and one person dies every 6 s due to diabetes [2]. In order to diagnose the disease, doctors use a blood glucose test. However, special care must be taken to prevent serious DM complications such as nephropathy (kidney), retinopathy (eye), cardiovascular, and neuropathy (nerve) diseases [3–5]. While one of the most unnoticed complications is cardiovascular autonomic neuropathy (CAN), a least recognized and understood complication of diabetes is diabetic autonomic neuropathy (DAN) [4]. CAN causes cardiac abnormalities; therefore monitoring of abnormalities using heart rate variability (HRV) can help to detect at an early stage [5]. There are various noninvasive techniques reported using fundus images to detect diabetes and diabetic complications [6–13]. These studies provide recognition using features from multiple regions [11,13] of fundus images [6,7], tongue [8] and face [9,10,12].

Recently, many methods have been implemented successfully using HRV signals to diagnose diabetes [14–20]. Acharya et al. [18] used discrete wavelet transform (DWT) features (of entropy and energy) obtained from the HRV signals and decision tree (DT) classifier to diagnose diabetes. Their method yielded an accuracy of 92.02%,

* Corresponding author.
E-mail address: oyildirim@munzur.edu.tr (O. Yildirim).

https://doi.org/10.1016/j.compbiomed.2019.103387
Received 4 June 2019; Received in revised form 8 August 2019; Accepted 8 August 2019
Available online 09 August 2019

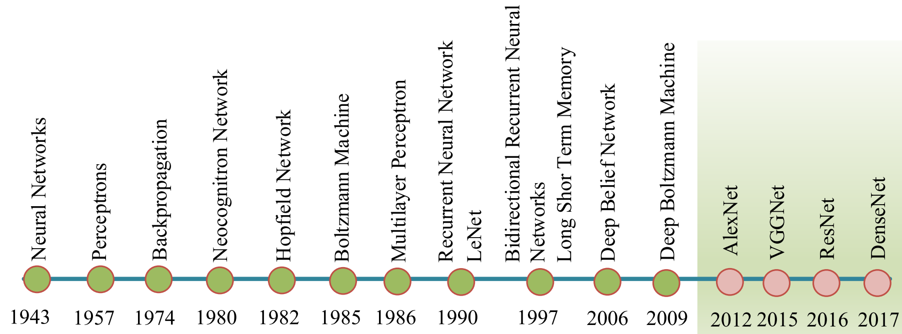

Fig. 1. Evolution of artificial intelligence (AI).

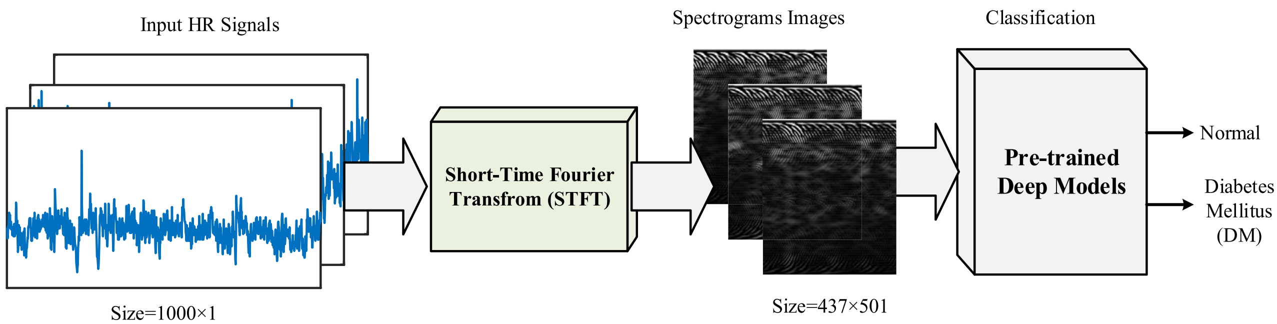

Fig. 2. A block representation of the proposed recognition system.

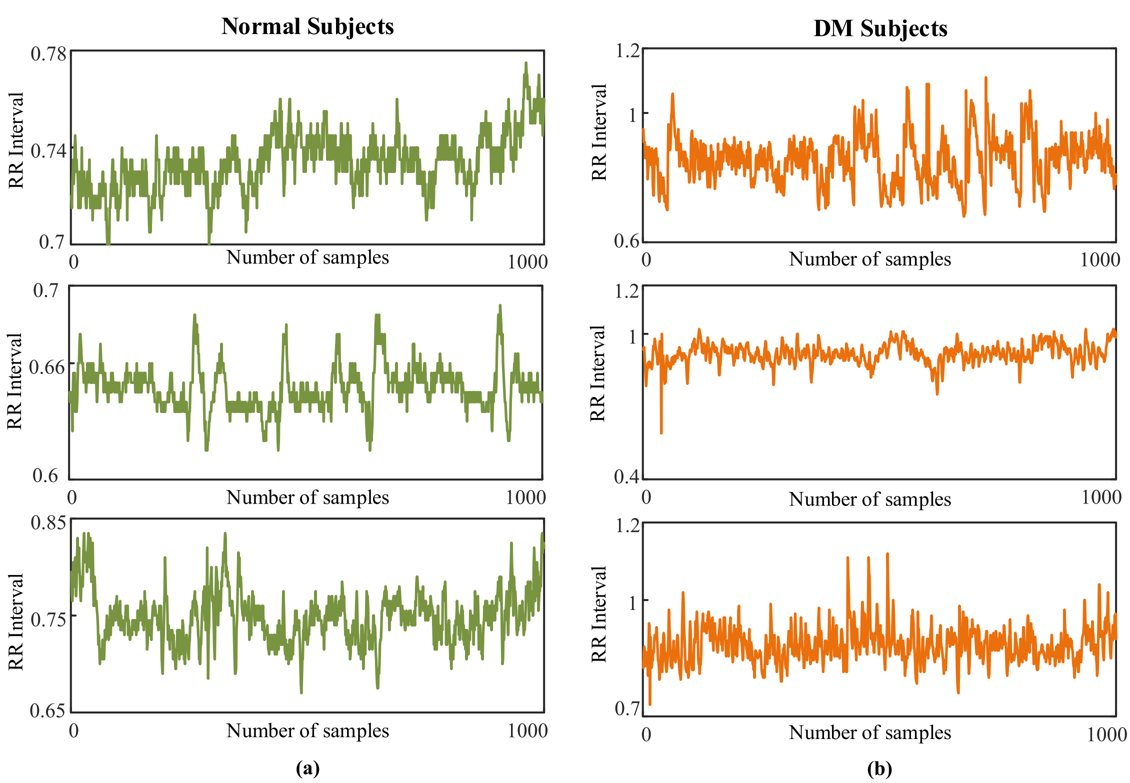

Fig. 3. Typical HR signal samples: a) Normal, and b) DM subjects.

sensitivity of 92.59%, and specificity of 91.46%. In another study [19], nonlinear features extracted from the HRV signals with the AdaBoost classifier obtained the highest average accuracy of 90%, sensitivity of 92.5%, and specificity of 88.7%. The same group [20] developed a novel diabetes index approach for diagnosis of diabetic neuropathy using features extracted from HRV signals. In Ref. [21], time, frequency, and nonlinear domain techniques were utilized to analyze normal and diabetic HR signals. They showed that nonlinear HRV

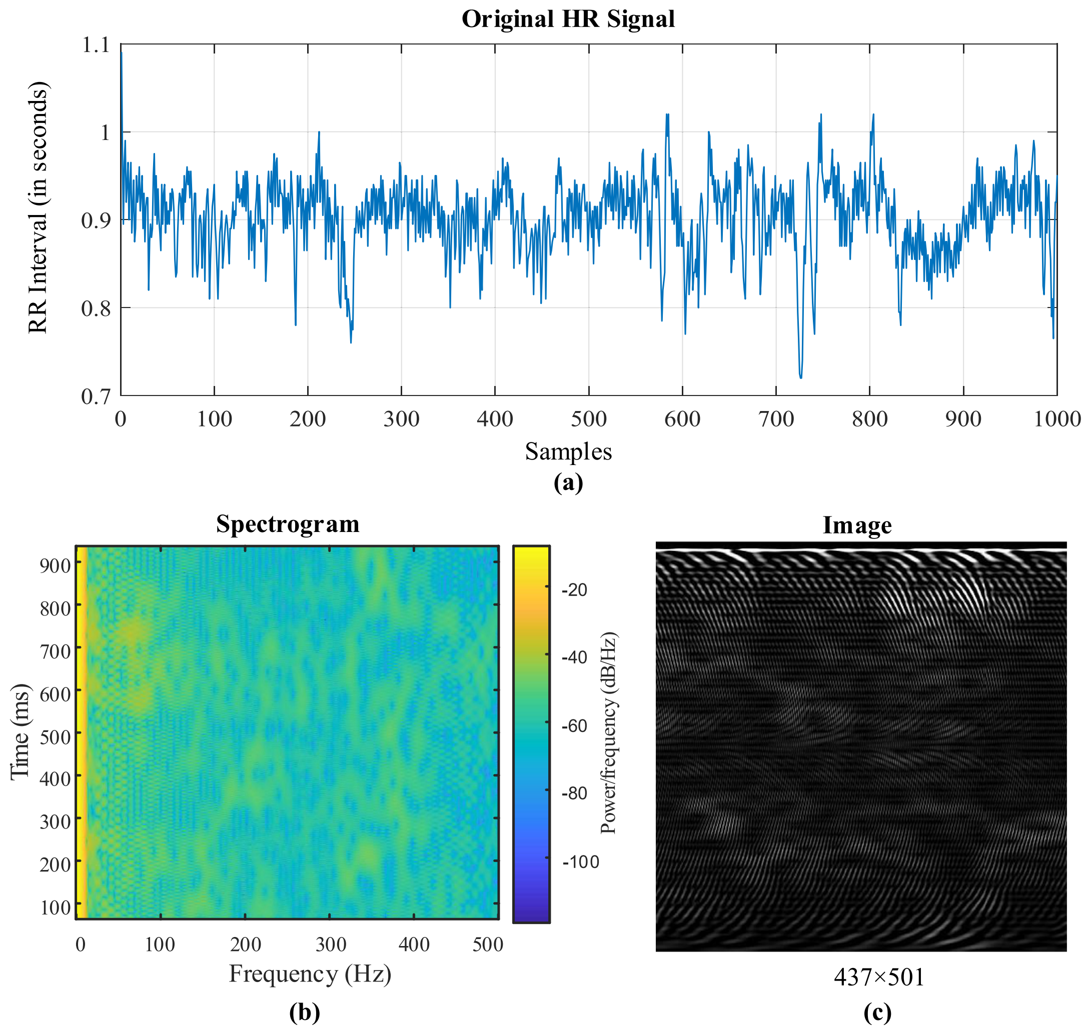

Fig. 4. Conversion of the original HR signals to frequency spectrum representative images: a) 1D HR signal, b) spectrogram of (a), and c) 2D grayscale image of (b).

analysis is more effective than time and frequency methods. Pachori et al. [22] classified diabetic and normal classes using features computed from intrinsic mode functions (IMFs) obtained from the empirical mode decomposition (EMD) of RR-interval signals. Swapna et al. [23] used the higher order spectra (HOS) method on HR signals. Their technique obtained the maximum accuracy of 90.5% using Gaussian mixture model (GMM). Using linear regression, Nolan et al. [24] performed a gender-based relationship analysis between HRV measures and the duration of type 2 diabetes. They reported gender-based distinctions among vagal-heart rate modulation, duration of diabetes, and total R-R variability in the HRV signals. Trunkvalterova et al. [25] employed the multiscale entropy (MSE) analysis to detect subtle abnormalities in young type 1 diabetes patients' cardiovascular system. Seyd et al. [26] applied frequency and time approaches to discern normals from DM patients by analyzing HR signals. For time and frequency domain analysis, they have used the ECG signals of 16 DM patients and 16 normal subjects. Mercaldo et al. [27] utilized different machine learning methods to differentiate diabetes affected patients from controls. Using the Hoeffding tree algorithm, they obtained a precision value of 77%.

The classical machine learning methods used to diagnose diabetes have difficulties. Feature extraction is one of the most important steps in traditional machine learning systems. The performance of the machine learning system depends on feature extraction. Extraction of the best performing features is done by the trial and error method, which is time-consuming. Deep learning performs automatic feature learning [28–30], and it mimics the structure of the human brain. The emergence of new approaches and powerful computational resources to compute and train the enormous amount of data have led to the rapid growth in the development of deep neural networks. Fig. 1 shows the evolution of artificial intelligence (AI). There are many applications of deep learning in biomedical image and signal processing studies [31–37]. Pratt et al. [38] have used a convolutional neural network (CNN) to classify diabetic retinopathy (DR) stages. Their network has reached a classification accuracy of 75% using 5 K validation images.

In this study, we used deep learning CNN models for the detection of diabetic subjects using HR signals. We have used the most popular deep learning pre-trained models: AlexNet, VggNet, DenseNet, and ResNet, trained using large image datasets, to achieve a higher detection performance. We transformed the HR signals into spectrogram images for pre-trained models. In this way, we achieved a significant improvement in classification performance. To the best of our knowledge, the proposed study is the first work to apply the 2-dimensional deep transfer learning approach using 1-dimensional HR signal data. The overall contribution of this study is summarized as follows:

- Provided an effective classification of DM subject with a complete end-to-end structure without requiring any hand-crafted feature extraction techniques.
- A deep learning-based approach has been developed using HR signals.
- Spectrogram images enabled pre-trained deep learning models to be

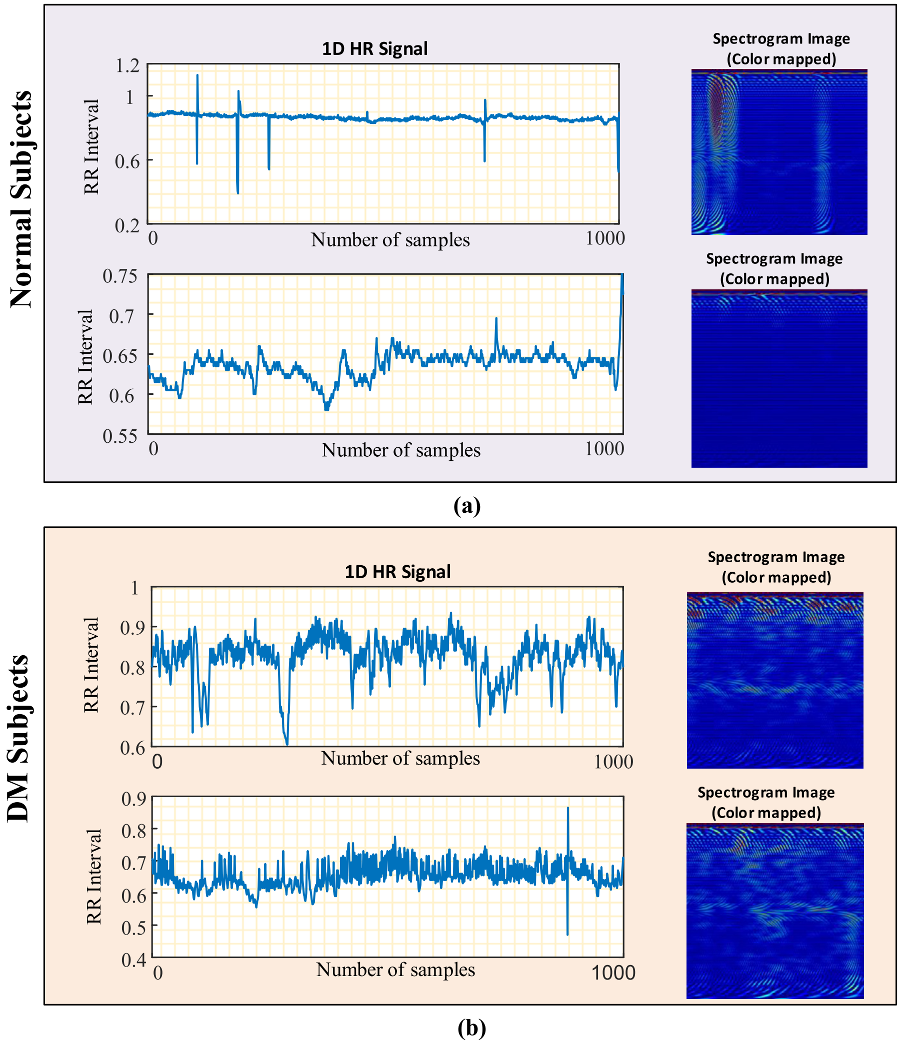

Fig. 5. Graphical representation of spectrogram images and respective HR signals: a) Normal subject b) DM subject.

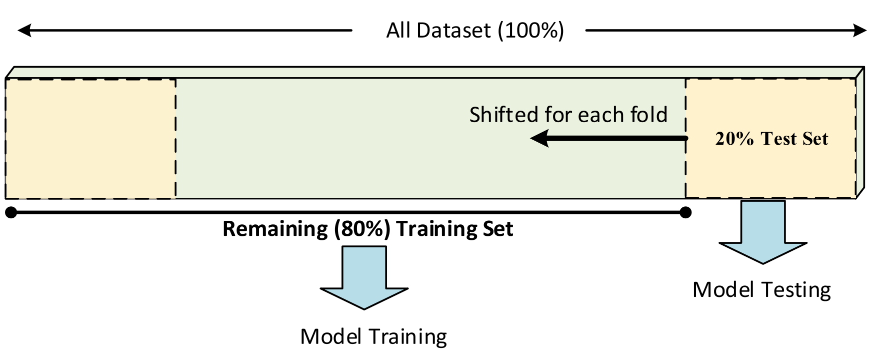

Fig. 6. Graphical representation of the test and training data sets for the 1D-CNN model.

trained on a small dataset (71 normal and 71 DM).
- Using deep transfer learning, the difficulties in the stages of model training and design is eliminated.

2. Material and methods

In this study, a deep learning framework is proposed for the detection of DM using HR signals. In order to benefit from the performance of pre-trained deep learning models which have been trained on the ImageNet database, HR signals are transformed into images having more visual representations. For this purpose, 1-dimensional signal data is converted to 2-dimensional gray images by the Short-Time Fourier Transform (STFT) method. The images having visual representations of the frequency spectra are used to train and test various popular pre-trained models. The block representation of the proposed method is shown in Fig. 2.

2.1. HR dataset

The ECG signals used in the study were obtained from 15 DM

Table 1
The layers and layer parameters of the 1D-CNN model designed for this work.

No	Layer Name	Filter Size	Kernel Size	Other Parameters	Output Size
0	Input Layer	–	–	–	1000 × 1
1	Conv-1D	64	5	Strides = 3	332 × 64
2	Conv-1D	64	3	Strides = 1	330 × 64
3	Max-Pooling			Strides = 2, Pool size = 2	165 × 64
4	Dropout	–	–	Rate = 0.1	165 × 64
5	Conv-1D	128	3	Strides = 1	163 × 128
6	Conv-1D	128	5	Strides = 1	159 × 128
7	Max-Pooling	–	–	Strides = 2, Pool size = 2	79 × 128
8	Conv-1D	256	2	Strides = 1	78 × 256
9	Conv-1D	256	3	Strides = 1	76 × 256
10	Max-Pooling	–	–	Strides = 2, Pool size = 2	38 × 256
11	Conv-1D	64	3	Strides = 1	36 × 64
12	Conv-1D	64	5	Strides = 1	32 × 64
13	Max-Pooling	–	–	Strides = 2, Pool size = 2	16 × 64
14	Flatten	–	–	–	1024
15	Dense (Relu)	–	–	Hidden units = 64	64
16	Dense (Softmax)	–	–	Hidden units = 2	2

*Hyper Parameters: Optimizer = Adam, Batch size = 8, Learning rate = 0.001

patients (7 female and 8 male) and 15 normal (5 female and 10 male) subjects for a 1 h duration [18]. In this study, diabetes patients with the disease for 5–15 years and age ranging from 50 to 70 years were considered. The subjects in the normal group were between 40 and 60 years old. The ECG signals were recorded at Kasturba Medical Hospital (KMH), Manipal, India. The pre-processing steps performed on the signals were as follows:

- Signals were recorded at 500 Hz sampling rate using the BIOPACTM system.
- 15 Hz cut-off frequency low-pass filter for noise removal and 0.3 Hz cut-off frequency high-pass filter to remove baseline wander were used.
- A band reject filter (50 Hz central frequency) was applied to clear away the power-line interfacing noise.
- The Pan-Tompkins algorithm [39] was used for RR point detection.

A total of 142 data files with each file containing 1000 samples were used. Hence, we have used 71 DM and 71 normal data files in this work. Fig. 3 shows the sample HR signals belonging to normal and DM subjects.

2.2. Spectrogram images

The Fourier transform (FT) is a crucial methodology which is not useful when the signal is nonstationary, and its spectral content is changing [40]. The HR signals represent nonstationary behavior by having altered frequencies and amplitudes over time. This change cannot be captured by FT based analysis. Therefore, we preferred to use a method which could capture these changes. The short-time Fourier transform (STFT) is a general-purpose FT based function, which can help the deep learning structure to extract hidden features from the spectrogram images effectively. STFT has been used in the past with CNN, and Long Short-Term Memory (LSTM) architectures for speech recognition [41] and motor imagery brain-computer interface recognition [42] tasks. The speech signal is a time-varying signal, and STFT can create a 2-dimensional representation of this signal. From the segmented audio data, STFT is used to generate binary images of speech and music [43]. Similarly, STFT can be used to generate a 2-dimensional representation of HR signals. In another signal classification study, the STFT and CNN combination has been successfully applied to EEG signals [44]. Salem et al. [45] have used spectrogram images of ECG signals for classification based on deep transfer learning. In our present study, we have converted HR signals to spectrogram images for use with pre-trained 2D CNN models. In Fig. 4, the flow of the conversion of the original HR signals to spectrogram images is provided.

HR signals with 1000 samples after passing through STFT are converted into spectrograms, which represent the frequency spectra. The spectrogram indicates changes in the frequency spectra of the original signal, which can be observed over time. The grayscale images (Fig. 4 (c)) of STFT images of size 437 × 501 pixels is presented. Fig. 5 shows the original HR signals of two normal and two DM subjects and the corresponding spectrogram images.

2.3. Deep networks

Deep learning is of great importance in many areas as well as in the medical field. Deep learning models, especially CNNs, have been successfully used in medical applications such as detection [46] and classification [47]. Krizhevsky et al. [48] ranked the first by successfully classifying images in the ImageNet large-scale visual recognition challenge (ILSVRC) with the developed deep learning-based CNN model (AlexNet) in 2012. In 2014, the VGGNet [49] CNN model demonstrated a better classification performance than AlexNet in ILSVRC. In the following year, the ResNet model developed by He et al. [50], which won the first place award for classification, detection, and segmentation tasks at ILSVRC 2015. The ResNet model was developed with the skip connection technique, which is deeper than the previous models (AlexNet and VggNet). The skip connections, also known as residual connections, attach every residual block to the next blocks. This technique enables information flow throughout the network. Therefore, it allows training of the CNN even with 1000 layers. Huang et al. [51] introduced densely connected convolutional networks (DenseNet) using a similar shortcut technique which connects every layer of the network to the following layers. DenseNet architecture generally showed high performance in classification problems such as ResNet architecture. However, DenseNet architecture used fewer parameters and required fewer computations than ResNet in training the model. Besides, one of the advantages of DenseNet model is that it exhibits better classification performance using small datasets [51].

The transfer learning method plays a vital role in finding a solution for classification problems. A CNN usually requires a large dataset to train with a large number of labeled data, and it requires higher computational power. One of the main problems in medical data analysis is the limited number of annotated data. The process of labeling data by experts is expensive and time-consuming. Therefore, instead of training the CNN model from scratch, we used the weights of the pre-trained models which already learned the distinguishing representations from a different but similar task. In other words, in the transfer learning technique, the weights of pre-trained models are transferred to the present model. In this way, small datasets are trained with a low computational cost.

3. Experimental results

In this study, HR signals are used for the detection of diabetes patients, and these signals are classified using deep learning based approaches. We have performed this in two steps. First, a CNN model is designed for raw HR signals, and the performance of this model is investigated. In the second step, we have tried to increase detection performance by examining the cases where the developed CNN model is inadequate. Therefore, the input signals are converted to image datasets. Then, the obtained images are classified with pre-trained models.

The training and testing of the proposed deep models are carried out on a Linux server with the Ubuntu 16.04 operating system using 11 GB of memory, including an NVIDIA GeForce GTX 1080 TI graphics card. The results are evaluated using the k-fold cross-validation (CV)

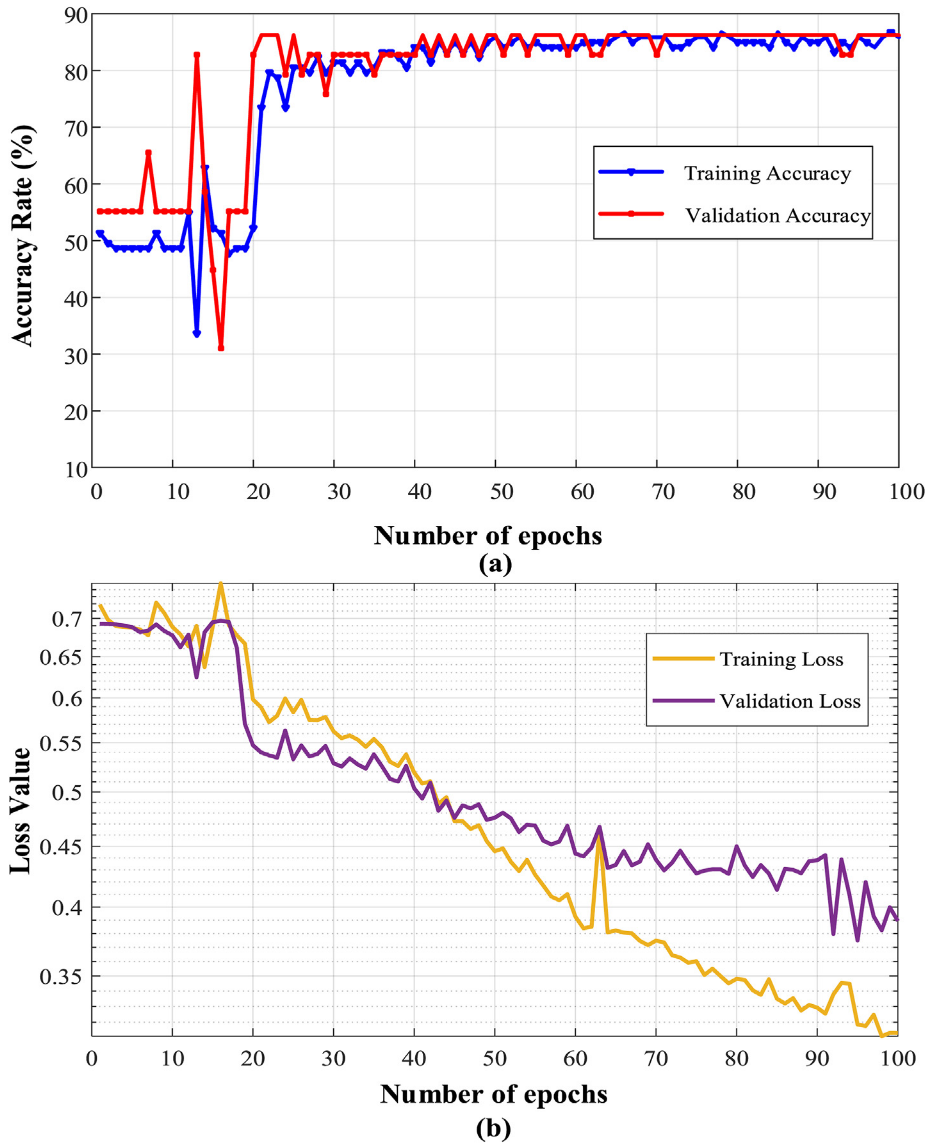

Fig. 7. The performance graphs of 1D CNN model obtained during 100 epochs: a) Accuracy, and b) Loss.

Table 2
1D CNN model's performance values for 5-fold test data.

Fold Number	Accuracy (%)	Sensitivity (%)	Precision (%)	F1-Score
Fold-1	86.21	100	69.23	81.82
Fold-2	79.31	85.71	75.00	80.00
Fold-3	86.21	84.21	94.12	88.89
Fold-4	89.66	100	80.00	88.89
Fold-5	89.66	92.86	86.67	89.66
Mean ± Std.	86.21 ± 4.22	92.55 ± 7.54	81.00 ± 9.74	85.85 ± 4.56

strategy. Thus, one of the folds is used as a validation set, and the remaining folds are used as the training set. The k value for HR signals is set to 5. Therefore, 20% of the data is reserved for testing, while the remaining 80% is used in the training phase. Fig. 6 shows the block representation of the training and testing dataset presented to the proposed models.

3.1. 1D CNN model

For the classification of HR signals, a 16-layer CNN model is constructed. The 1000 × 1 HR signals are input to this model. The CNN model is composed of convolution, pooling, and dense layers, and includes a dropout layer to prevent overfitting. Table 1 presents the layer parameters of the proposed model. The model parameters are adjusted by the brute-force technique, similar to our previous deep learning works done using ECG and EEG signal data [31,33,46,47]. The Keras deep learning library is used to construct, train, and test the model.

In this work, 80% of the HR signals were used for training, and the remaining 20% for validation. The performance evaluations were made using a 5-fold CV strategy. The training and validation accuracy and loss graphs of the model for a fold for 100 epochs are given in Fig. 7.

The CNN model completed the training process without having any overfitting problem on one dimensional HR signals. After training, the training accuracy reached 86.73%, while the validation accuracy remained at 86.21%. Table 2 presents the performance values of the CNN network using a 5-fold CV strategy. The CNN network can reach 86.21% ± 4.22 average accuracy with 5-fold CV. In the end, we concluded that the 1-dimensional CNN model did not reach the desired level of success.

3.2. Pre-trained 2D CNN models

The most critical stage of the experimental studies is converting HR

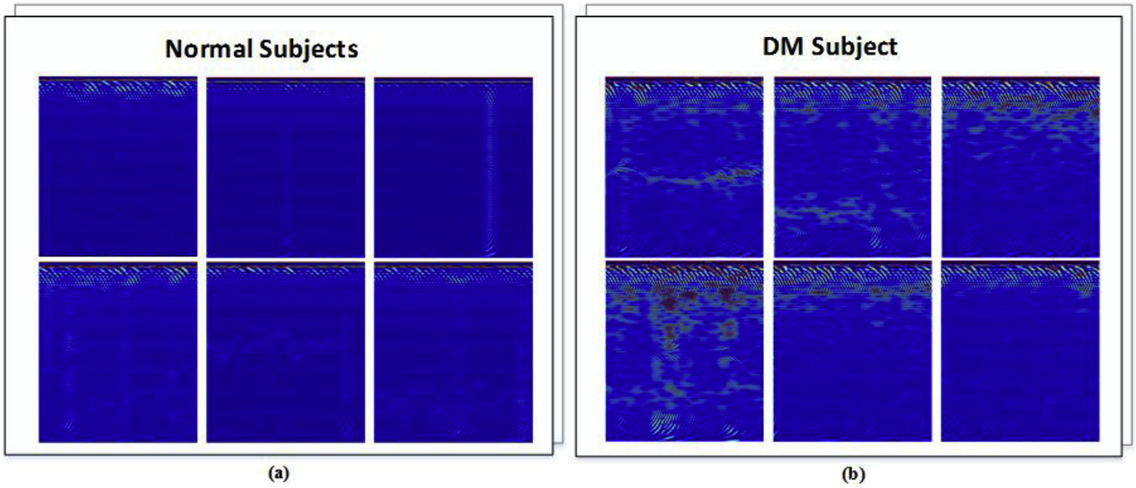

Fig. 8. Sample images obtained after converting HR signals to frequency spectrum images: a) Normal subject, and b) DM subject.

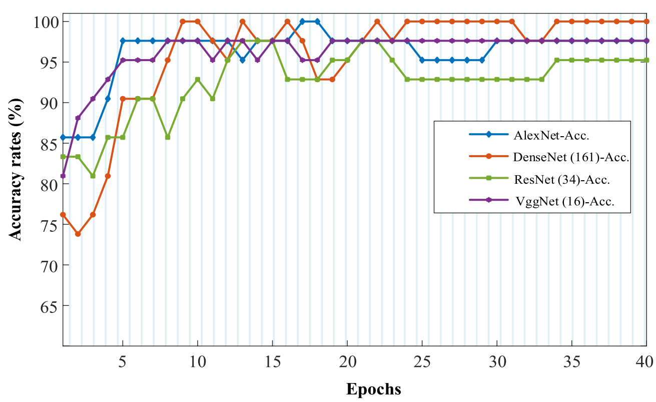

Fig. 9. Graphs of accuracies obtained for different epochs with various pre-trained models: AlexNet (blue line), DenseNet (orange line), ResNet (green line) and VggNet (purple line). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

signals into images that can be processed with pre-trained 2D-CNN models to improve classification performance. Popular pre-trained models such as AlexNet, VggNet, DenseNet, and ResNet are trained and tested on these image data. For this purpose, the STFT method is used on 1-dimensional HR signals to obtain 2-D images, which indicate the visual representations of frequency spectra. Fig. 8 shows a few images that represent frequency spectra corresponding to normal and DM subjects.

It is difficult to discriminate the two classes (normal/DM) using spectrogram images visually. Deep learning methods can perform the recognition process with high performance by extracting abstract features from these images. The performance of the networks can be significantly increased, especially with the use of pre-trained models which are trained on large data. In our study, we used AlexNet, DenseNet, ResNet, and VggNet pre-trained models, which performed well for image classification tasks in the field of deep learning. We have trained only the last layers (fully-connected) of pre-trained models. Fig. 9 shows the training performance of pre-trained models using the same-fold data.

The proposed pre-trained models are trained for 40 epochs. When the performances of the pre-trained models are evaluated, we witnessed that the DenseNet pre-trained model obtained the best classification accuracy among others. While the performance of AlexNet and VggNet are similar, the lowest performance is achieved with ResNet. The sensitivity, specificity, precision, F1-score, and accuracy measurements are used to evaluate the performance of the pre-trained models. The performance values of the proposed pre-trained models for the 5-fold cross-validation (CV) strategy are given in Table 3.

The classification performances of the AlexNet and ResNet-34 models are interestingly similar. The average accuracy values of both networks showed the lowest performance of 95.7% ± 1.0 The highest average accuracy of 97.6% ± 2.3 is obtained using DenseNet-161. In addition, the sensitivity value of DenseNet-161 is increased to as high as 100% with the 5-fold CV strategy. Fig. 10 shows the accuracy and loss graphs of the DenseNet-161 model during 40 epoch periods for a fold.

It can be observed from the accuracy graph (Fig. 10(a)) that DenseNet-161 model reached the highest classification accuracy of 100% at the end of 40 epochs for the test images. When the loss graphs are analyzed (Fig. 10(b)), both training and validation loss values decreased during the training of the model. It may be noted that, when the training of the DenseNet-161 model is completed, the validation loss is still less than the training loss (underfitting). The main reason for this is that the DenseNet-161 has correctly classified each image in the validation set; therefore, the classification accuracy of the model has yielded 100%, and the validation loss reached the minimum value. It can be observed from this result that the model's predictions on the images in the validation set, which it has not seen before during training, were very successful. This outcome assured that the training of the model was at the desired level. Fig. 11 shows the confusion matrix acquired with the 5-fold cross validation (CV) strategy using the DenseNet-161 model with diabetes test data.

The model misclassified only five images in the DM class, and

Table 3
The performance values obtained by various pre-trained models with 5-fold test data (mean ± standard deviation).

Models	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	Accuracy (%)
AlexNet	93.66 ± 3.6	96.60 ± 0.1	92.13 ± 0.7	92.87 ± 1.9	95.72 ± 1.0
VggNet (16)	95.26 ± 4.3	97.29 ± 2.8	94.13 ± 5.7	94.51 ± 2.0	96.67 ± 1.3
ResNet (34)	93.80 ± 3.5	96.59 ± 0.1	92.13 ± 0.7	92.93 ± 1.6	95.72 ± 1.0
DenseNet (161)	100	96.72 ± 3.3	92.33 ± 7.6	95.88 ± 4.1	97.62 ± 2.3

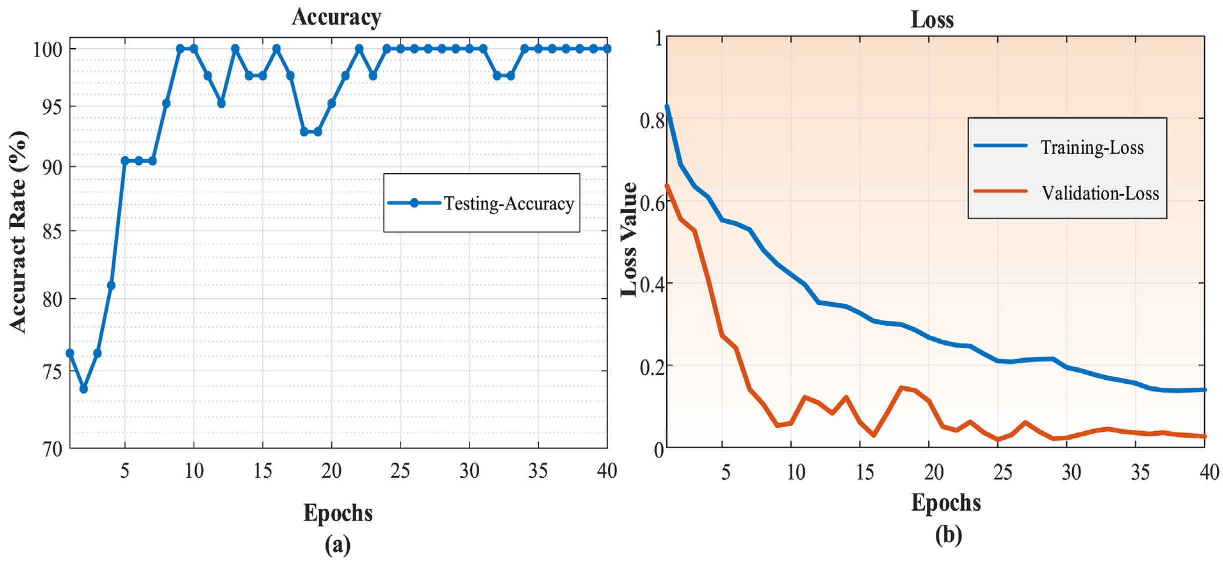

Fig. 10. The performance graphs of DenseNet-161 model on the diabetes test dataset: a) testing accuracy b) training and validation loss.

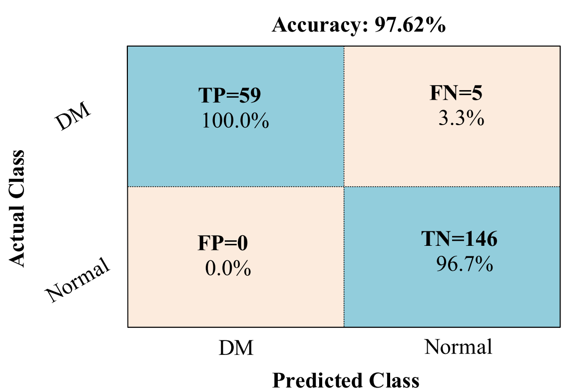

Fig. 11. Overall confusion matrix obtained with 5-fold for the DenseNet-161 model.

correctly classified all normal class images in all folds. The 5-fold average classification accuracy value is 97.6% ± 2.3. The performance of the DenseNet-161 model on each fold data is given in Table 4.

The developed 16-layer CNN model has been trained from scratch. Initially, the weights of the proposed network were set randomly and then adjusted through backpropagation. The whole network is trained until optimal performance is reached. However, for our 2-dimensional data, we have used pre-trained models (AlexNet, VggNet, ResNet, and DenseNet) to detect a diabetic subject using frequency spectrum images. These pre-trained models have already been trained to perform other classification tasks. We have transferred the weights of the proposed pre-trained models. During the training of pre-trained CNN models, we have only trained the fully connected layers of the pre-trained models. Therefore, the deep learning model used for 1-dimensional data trained from scratch requires more epochs (100) than the pre-trained models.

Our findings show that the results obtained using pre-trained models have yielded better classification accuracy using 2-dimensional data than the proposed 16-layer deep learning model with 1-dimensional data. The AlexNet, VggNet, ResNet, and DenseNet pre-trained CNN models have been trained with more than a million images belonging to 1000 different categories. These CNN models have already learned valuable representations from various images. We have used this gained knowledge by implementing the transfer learning technique for 2-dimensional data. For 1-dimensional data, the proposed deep learning model only learned the representations from the given data. Therefore, the classification accuracy of the constructed model is lower than for the pre-trained CNN models. In the future, we hope that such transfer learning techniques can also be employed for 1-dimensional data to obtain high performance.

4. Discussion

There are a few remarkable state-of-the-art studies performed for the automated detection of DM subjects using HR signals. In these studies, the discrete wavelet transform (DWT) [18], empirical mode decomposition (EMD) [22], higher order spectra (HOS) [23], non-linear analysis [21], statistical methods [20] have been employed for feature extraction. For automated detection of DM, obtained features are input to the shallow-structured classifiers, such as decision tree (DT) [18], AdaBoost [19,20], Gaussian mixture model (GMM) [23], and support vector machines (SVM) [49]. A few other signal processing methods, coupled with machine learning models, have been used to detect DM automatically [21,22]. The accuracies obtained using these studies have a range between 90% and 92%. Our present study employed a deep learning-based approach for DM detection. With this completely end-to-end structure, the signals are classified without requiring any handcrafted feature extraction. Table 5 shows a comparison of

Table 4
The performances of DenseNet-161 model for each of 5-fold DM classification.

Fold Number	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Fold-1	100	100	100	100	100
Fold-2	100	96.77	91.67	95.65	97.62
Fold-3	100	93.10	86.67	92.86	95.24
Fold-4	100	93.75	83.33	90.91	95.24
Fold-5	100	100	100	100	100

Table 5
Comparison of performances for automated detection of diabetes using the same HRV signal database.

Study	Methods	Results/Findings
Acharya et al. [18]	DWT & Decision Tree	Sen = 92.59%, Acc = 92.02%
Acharya et al. [19]	Nine nonlinear measures & AdaBoost	Sen = 92.50%, Acc = 90.00%
Acharya et al. [20]	Diabetic integrated index & AdaBoost	Sen = 87.50%, Acc = 86.00%
Faust et al. [21]	Six non-linear features & student t-test	Non-linear analysis is more effective than frequency and time domain analysis
Pachori et al. [22]	EMD & Kruskal–Wallis statistical test	Significant difference between diabetic and normal classes ($p < 0.05$)
Swapna et al. [23]	HOS & GMM	Sen = 85.70%, Acc = 90.50%
Jian and Lim [52]	HOS & SVM	Sen = 70.97%, Acc = 79.93%
This study	End-to-end 1D HR signals & 1D-CNN model	Sen = 92.55% ± 7.5, Acc = 86.21% ± 4.2
This study	End-to-end 2D frequency spectrum images & 2D-CNN DenseNet161	Sen = 100%, Acc = 97.62% ± 2.3

performance for automated detection of diabetes using the same HRV signal database. The proposed study achieved a 97.62% ± 2.3 average accuracy and a sensitivity of 100%, which outperformed existing state-of-the-art studies in the literature.

The advantage of the present study is the transfer of weights from popular models, which trained on two-dimensional large image data, to a small number of signal data. Thus, the constraint in the construction and training phase of deep models is eliminated. Also, such a developed model (in our present work) has yielded higher classification performance. In the future, we can use this model to detect diabetes at an early stage, and it also can be used to detect other diseases from biomedical signals. The main disadvantage of this work is that we have used a small database. Deep learning models consist of many layers with millions of parameters in these layers. The models process the data iteratively to obtain optimal parameters during the training phase. Repeated processing of data results in various problems with small data because the model memorizes the training data and fails to recognize test data, which it has not seen (overfitting problem). This overfitting problem can be addressed using a transfer learning technique by employing the models that have been previously trained on large datasets. With the proposed method, it is possible to optimize the existing parameters (learned features) instead of setting all model parameters to zero.

5. Conclusion

In this study, we proposed a deep transfer learning based approach using spectrogram images obtained from HR signals to detect diabetes patients automatically. AlexNet, VggNet, ResNet, and DenseNet CNN pre-trained models trained on 2D image data were used for the evaluation of one-dimensional HR signals. A total of 142 segments (71 normal and 71 DM) obtained from 30 subjects (15 normal and 15 DM) were used in this study. The DenseNet-161 CNN model achieved a 97.62% ± 2.3 accuracy and a 100% sensitivity performance with the 5-fold CV strategy. Hence, diabetes patients can be discerned accurately using the pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals. In the future, we intend to improve model accuracy by training it with more data and also focus on the early detection of diabetes. We also plan to explore the possibility of using more HR signals to evaluate the performance of our developed model.

Conflicts of interest

All authors declare that there is no conflict of interest in this work.

References

[1] I. Benjamin, R.C. Griggs, E.J. Wing, J.G. Fitz, Andreoli and Carpenter's Cecil Essentials of Medicine E-Book, Elsevier Health Sciences, 2015.
[2] International Diabetes Federation, IDF Diabetes Atlas, eighth ed., (2017) http://www.diabetesatlas.org/resources/2017-atlas.html.
[3] W.T. Cade, Diabetes-related microvascular and macrovascular diseases in the physical therapy setting, Phys. Ther. 88 (11) (2008) 1322–1335.
[4] A.I. Vinik, R.E. Maser, B.D. Mitchell, R. Freeman, Diabetic autonomic neuropathy, Diabetes Care 26 (5) (2003) 1553–1579.
[5] A.I. Vinik, D. Ziegler, Diabetic cardiovascular autonomic neuropathy, Circulation 115 (3) (2007) 387–397.
[6] M. Niemeijer, et al., Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs, IEEE Trans. Med. Imaging 29 (1) (2009) 185–195.
[7] U.R. Acharya, C.M. Lim, E.Y.K. Ng, C. Chee, T. Tamura, Computer-based detection of diabetes retinopathy stages using digital fundus images, Proc. Inst. Mech. Eng. H J. Eng. Med. 223 (5) (2009) 545–553.
[8] B. Zhang, B.V. Kumar, D. Zhang, Detecting diabetes mellitus and nonproliferative diabetic retinopathy using tongue color, texture, and geometry features, IEEE Trans. Biomed. Eng. 61 (2) (2013) 491–501.
[9] B. Zhang, D. Zhang, Noninvasive diabetes mellitus detection using facial block color with a sparse representation classifier, IEEE Trans. Biomed. Eng. 61 (4) (2013) 1027–1033.
[10] T. Shu, B. Zhang, Y.Y. Tang, An extensive analysis of various texture feature extractors to detect Diabetes Mellitus using facial specific regions, Comput. Biol. Med. 83 (2017) 69–83.
[11] J. Li, D. Zhang, Y. Li, J. Wu, B. Zhang, Joint similar and specific learning for diabetes mellitus and impaired glucose regulation detection, Inf. Sci. 384 (2017) 191–204.
[12] T. Shu, B. Zhang, Y.Y. Tang, An improved noninvasive method to detect Diabetes Mellitus using the Probabilistic Collaborative Representation based Classifier, Inf. Sci. 467 (2018) 477–488.
[13] J. Li, B. Zhang, G. Lu, J. You, D. Zhang, Body surface feature-based multi-modal learning for diabetes mellitus detection, Inf. Sci. 472 (2019) 1–14.
[14] U.R. Acharya, H. Fujita, V.K. Sudarshan, V.S. Sree, L.W.J. Eugene, D.N. Ghista, R. San Tan, An integrated index for detection of sudden cardiac death using discrete wavelet transform and nonlinear features, Knowl. Based Syst. 83 (2015) 149–158.
[15] Y. Cheng, F. Wang, P. Zhang, J. Hu, Risk prediction with electronic health records: a deep learning approach, Proc. 2016 SIAM International Conference on Data Mining, 2016, pp. 432–440.
[16] M.R.K. Mookia, U.R. Acharya, C.K. Chua, C.M. Lim, E.Y.K. Ng, A. Laude, Computer-aided diagnosis of diabetic retinopathy: a review, Comput. Biol. Med. 43 (12) (2013) 2136–2155.
[17] M.R.K. Mookiah, U.R. Acharya, R.J. Martis, C.K. Chua, C.M. Lim, E.Y.K. Ng, A. Laude, Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: a hybrid feature extraction approach, Knowl. Based Syst. 39 (2013) 9–22.
[18] U.R. Acharya, K.S. Vidya, D.N. Ghista, W.J.E. Lim, F. Molinari, M. Sankaranarayanan, Computer-aided diagnosis of diabetic subjects by heart rate variability signals using discrete wavelet transform method, Knowl. Based Syst. 81 (2015) 56–64.
[19] U.R. Acharya, O. Faust, N.A. Kadri, J.S. Suri, W. Yu, Automated identification of normal and diabetes heart rate signals using nonlinear measures, Comput. Biol. Med. 43 (10) (2013) 1523–1529 2013.
[20] U.R. Acharya, O. Faust, S.V. Sree, D.N. Ghista, S. Dua, P. Joseph, V.I.T. Ahamed, N. Janarthanan, T. Tamura, An integrated diabetic index using heart rate variability signal features for diagnosis of diabetes, Comput. Methods Biomech. Biomed. Eng. 16 (2) (2013) 222–234.
[21] O. Faust, U.R. Acharya, F. Molinari, S. Chattopadhyay, T. Tamura, Linear and non-linear analysis of cardiac health in diabetic subjects, Biomed. Signal Process. Control 7 (3) (2012) 295–302.
[22] R.B. Pachori, P. Avinash, K. Shashank, R. Sharma, U.R. Acharya, Application of empirical mode decomposition for analysis of normal and diabetic RR-interval signals, Expert Syst. Appl. 42 (9) (2015) 4567–4581.
[23] G. Swapna, U.R. Acharya, S. VinithaSree, J.S. Suri, Automated detection of diabetes using higher order spectral features extracted from heart rate signals, Intell. Data Anal. 17 (2) (2013) 309–326.
[24] R.P. Nolan, S.M. Barry-Bianchi, A.E. Mechetiuc, M.H. Chen, Sex-based differences in the association between duration of type 2 diabetes and heart rate variability, Diabetes Vasc. Dis. Res. 6 (4) (2009) 276–282.

[25] Z. Trunkvalterova, M. Javorka, I. Tonhajzerova, J. Javorkova, Z. Lazarova, K. Javorka, M. Baumert, Reduced short-term complexity of heart rate and blood pressure dynamics in patients with diabetes mellitus type 1: multiscale entropy analysis, Physiol. Meas. 29 (7) (2008) 817.
[26] P.A. Seyd, V.T. Ahamed, J. Jacob, P. Joseph, Time and frequency domain analysis of heart rate variability and their correlations in diabetes mellitus, Int. J. Biol. Life Sci. 4 (1) (2008) 24–27.
[27] F. Mercaldo, V. Nardone, A. Santone, Diabetes mellitus affected patients classification and diagnosis through machine learning techniques, Procedia comput. sci. 112 (2017) 2519–2528.
[28] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436.
[29] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT press, 2016.
[30] M. Coşkun, Ö. Yildirim, A. Uçar, Y. Demir, An overview of popular deep learning methods, Eur J Tech 7 (2) (2017) 165–176.
[31] O. Yildirim, U.B. Baloglu, U.R. Acharya, A deep learning model for automated sleep stages classification using psg signals, Int. J. Environ. Res. Public Health 16 (4) (2019) 599.
[32] S.K. Pandey, R.R. Janghel, Recent deep learning techniques, challenges and its applications for medical healthcare system: a review, Neural Process. Lett. (2019) 1–29.
[33] U.B. Baloglu, M. Talo, O. Yildirim, R. San Tan, U.R. Acharya, Classification of myocardial infarction with multi-lead ECG signals and deep CNN, Pattern Recognit. Lett. 122 (2019) 23–30.
[34] S.L. Oh, E.Y. Ng, R. San Tan, U.R. Acharya, Automated beat-wise arrhythmia diagnosis using modified U-net on extended electrocardiographic recordings with heterogeneous arrhythmia types, Comput. Biol. Med. 105 (2019) 92–101.
[35] O. Faust, Y. Hagiwara, T.J. Hong, O.S. Lih, U.R. Acharya, Deep learning for healthcare applications based on physiological signals: a review, Comput. Methods Progr. Biomed. 161 (2018) 1–13.
[36] N. Michielli, U.R. Acharya, F. Molinari, Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals, Comput. Biol. Med. 106 (2019) 71–81.
[37] P. Mamoshina, A. Vieira, E. Putin, A. Zhavoronkov, Applications of deep learning in biomedicine, Mol. Pharm. 13 (5) (2016) 1445–1454.
[38] H. Pratt, F. Coenen, D.M. Broadbent, S.P. Harding, Y. Zheng, Convolutional neural networks for diabetic retinopathy, Procedia comput. sci. 90 (2016) 200–205.
[39] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, IEEE Trans. Biomed. Eng. 32 (3) (1985) 230–236.
[40] S. Nisar, O.U. Khan, M. Tariq, An efficient adaptive window size selection method for improving spectrogram visualization, Comput. Intell. Neurosci. 2016 (2016).
[41] W. Lim, D. Jang, T. Lee, Speech emotion recognition using convolutional and recurrent neural networks, 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2016, pp. 1–4.
[42] Z. Wang, L. Cao, Z. Zhang, X. Gong, Y. Sun, H. Wang, Short time Fourier transformation and deep neural networks for motor imagery brain computer interface recognition, Concurrency Comput. Pract. Ex. 30 (23) (2018) e4413.
[43] P. Neammalai, S. Phimoltares, C. Lursinsap, Speech and music classification using hybrid form of spectrogram and fourier transformation, Signal and Information Processing Association Annual Summit and Conference(APSIPA), 2014, pp. 1–6. Asia-Pacific 2014.
[44] Y. Lu, H. Jiang, W. Liu, Classification of eeg signal by stft-cnn framework: identification of right-/left-hand motor imagination in bci systems, The 7th International Conference on Computer Engineering and Networks, vol 299, 2017, p. 1.
[45] M. Salem, S. Taheri, J.S. Yuan, ECG arrhythmia classification using transfer learning from 2-dimensional deep CNN features, 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS), 2018, pp. 1–4.
[46] Ö. Yıldırım, P. Pławiak, R.S. Tan, U.R. Acharya, Arrhythmia detection using deep convolutional neural network with long duration ECG signals, Comput. Biol. Med. 102 (2018) 411–420.
[47] M. Talo, U.B. Baloglu, Ö. Yıldırım, U.R. Acharya, Application of deep transfer learning for automated brain abnormality classification using MR images, Cogn. Syst. Res. 54 (2019) 176–188.
[48] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[49] Simonyan K, Zisserman A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint 2014; arXiv:1409.1556.
[50] 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, 2016, pp. 770–778.
[51] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
[52] L.W. Jian, T.C. Lim, Automated detection of diabetes by means of higher order spectral features obtained from heart rate signals, J. Med. Imag. Health Inf. 3 (3) (2013) 440–447.