



International Journal of Medical Engineering and Informatics

ISSN online: 1755-0661 - ISSN print: 1755-0653
<https://www.inderscience.com/ijmei>

Empirical wavelet decomposition of photoplethysmographic signal for hypertension risk stratification and detection of diabetes mellitus using machine learning techniques

Muzaffar Khan, Bikesh Kumar Singh, Neelamshobha Nirala

DOI: [10.1504/IJMEI.2022.10050838](https://doi.org/10.1504/IJMEI.2022.10050838)

Article History:

Received:	02 February 2022
Accepted:	15 June 2022
Published online:	12 December 2024

Empirical wavelet decomposition of photoplethysmographic signal for hypertension risk stratification and detection of diabetes mellitus using machine learning techniques

Muzaffar Khan*

Department of Biomedical Engineering,
National Institute of Technology,
Raipur, 492010, India
Email: mkhan.phd2016.bme@nitrr.ac.in
and

Department of Electronics and Telecommunication,
Anjuman College of Engineering and Technology,
Nagpur, India
Email: muzaffarkhanacet@gmail.com

*Corresponding author

**Bikesh Kumar Singh and
Neelamshobha Nirala**

Department of Biomedical Engineering,
National Institute of Technology,
Raipur, 492010, India
Email: bsingh.bme@nitrr.ac.in
Email: neelanir.bme@nitrr.ac.in

Abstract: Hypertension (HT) is a leading risk factor for cardiovascular disease (CVD), and the overlap of diabetes mellitus (DM) with hypertension can lead to severe complications. Presently, the diagnostic method for detecting hypertension and DM is unsuitable for large-scale screening. The proposed model uses a statistical feature extracted by decomposing the PPG signal into a sub-band signal using empirical wavelet transform (EWT), a comparative study conducted between various soft and hard computing classification models. The highest accuracy achieved by sequential neural network for the three categories, namely normal (NT) vs. prehypertension (PHT), NT vs. hypertension type 1 (HT-I), NT vs. hypertension type 2 (HT-II) in terms of F1 scores is 78.9%, 91.2% and 94%, respectively and F1 score of 97.9% for detection of DM-II patients. We conclude that soft computing techniques such as deep learning neural networks have shown superior performance compared to hard computing techniques. Furthermore, features selected using a hybrid feature selection technique were found to improve the classifier's performance. The main advantage of the proposed model that uses a decomposition technique is found to be more immune to noisy PPG signals, overcoming the limitation of the morphological-based model.

Keywords: hypertension; diabetes mellitus; photoplethysmographic; empirical wavelet transform; Hilbert transform ensemble classifier; deep learning neural network.

Reference to this paper should be made as follows: Khan, M., Singh, B.K. and Nirala, N. (2025) ‘Empirical wavelet decomposition of photoplethysmographic signal for hypertension risk stratification and detection of diabetes mellitus using machine learning techniques’, *Int. J. Medical Engineering and Informatics*, Vol. 17, No. 1, pp.74–88.

Biographical notes: Muzaffar Khan has 23 years of teaching experience and worked as a reviewer for reputed international journals. His research interest includes applications of machine learning and artificial intelligence in data analysis and cognitive learning. He has filed two patents AI for efficient assessment and prediction of human performance in collaborative learning environments and blockchain technology for detection of fake drugs in the pharmaceutical industry.

Bikesh Kumar Singh is working as an Assistant Professor in the Department of Biomedical Engineering at the National Institute of Technology Raipur, Raipur (Chhattisgarh) India. He is a member of International Professional Societies such as IEEE (senior member) and of many national professional bodies like CSI India, IETE India, ISCA India and IEI India. His research interest includes applications of machine learning and artificial intelligence in medical image analysis, biomedical signal analysis, computer-aided diagnosis, computer vision and cognitive science.

Neelamshobha Nirala is working as an Assistant Professor in the Department of Biomedical Engineering at the National Institute of Technology Raipur, Raipur (Chhattisgarh) India. He received his PhD in Biomedical Engineering from the National Institute of Technology Raipur, Raipur (Chhattisgarh) India. Her research interest includes applications of machine learning and artificial intelligence in the non-invasive diagnosis of diabetes mellitus, stenosis and peripheral arterial disease.

1 Introduction

Hypertension (HT) is commonly known as high blood pressure (BP) and is classified in to four categories based on the value of systolic BP (SBP) and diastolic BP (DBP). The BP values based on SBP and DBP value is categories into normal ($SBP < 120$ mmHg and $DBP < 80$ mmHg), prehypertension ($SBP 120\text{--}129$ mmHg and $DBP < 80$ mmHg), hypertension type I ($SBP 130\text{--}139$ mmHg) and type II ($SBP \geq 140$ mmHg and $DBP \geq 90$ mmHg) (Bundy et al., 2019). Hypertensive patients with diabetic mellitus (DM) overlap can have higher chances of increased peripheral resistance, which increases cardiovascular disease (CVD) risk factors (Petrie et al., 2017). The severity and widespread of HT and DM can be known from a report published in 2019 by the World Health Organization (WHO). The report estimated that 1.13 billion people suffer from hypertension worldwide, two-thirds of which are from low-income and middle-income countries. WHO protocol emphasises early diagnosis and continuous monitoring of BP and blood sugar values. The limitation of the current method of BP monitoring is cuff based, which is not suitable for large-scale screening and continuous monitoring.

Detection of DM requires a pathological laboratory-based test which is invasive and requires significant time. Patients suffering from DM experience fluctuating blood sugar values which require frequent monitoring. Recent research on optical-based technology

such as photoplethysmogram (PPG) resulted in devices such as pulse oximeters and heart rate meters that proved useful for large-scale screening, easy to operate, low cost, and reduced diagnostic time. Currently, significant research is directed toward developing cuffless Blood pressure measurement and DM detection devices using PPG. PPG technology offers a non-invasive diagnosis technology. It is based on the electro-optical technique that measures changes in blood volume in a peripheral vascular system that occurs during the cardiac cycle. The PPG signal system comprises an infrared light source and an infrared detection sensor. The system works on the principle of absorption and reflection of the near-infrared light source by blood and tissue is detected by optical sensors. The applications of a PPG signal in medical healthcare, which include hypertension and diabetes detection, can be found in the literature (Castaneda et al., 2018; Elgendi et al., 2019). Most current research uses time-domain analysis methods in which features are extracted from the morphological characteristics of the PPG signals, commonly employed methods for cuff-less BP measurement and DM detection. The diagnostic procedure is based on identifying morphological changes in the signal associated with the disease. The limitation of an expert diagnostic system based on time-domain analysis, the accuracy of a system, depends on the quality of a filtered stationary signal. Our proposed expert diagnosis system overcomes this limitation using a frequency domain analysis. Frequency domain analysis offers insight into non-stationary PPG signals, such as information about the hemodynamics of arterial tree, turbulence in the flow information is embedded in the signal's spectrum, and immune to noisy PPG signals (Akhtar et al., 2018). The proposed system uses a hybrid feature selection model based on majority voting in which features are selected based on voting scores.

Further, mutual information (MI) combined with the meta-heuristic such as a genetic algorithm (GA), is used to find global minima and avoid getting trapped into local minima. The advantage of using hybrid feature selection is that it takes the best of two and overcomes the limitation of an individual algorithm (Wang et al., 2014; Keerthika et al., 2019). Our proposed expert system uses advanced classifiers such as gradient boosting (GB), k-nearest neighbour (KNN) classifier, and sequential neural network. The main contribution of our work is that the proposed expert diagnostic system is more immune to noisy signals and provides a dual function of detecting hypertension and DM-II subject. In addition, this study provides insight into the intrinsic relationship between hypertension and DM-II patients.

The paper is organised as follows. Section 2 explains the material and methods used in this study. In contrast, Section 3 describes the feature extraction and selection, followed by results and discussion in Section 4 and conclusions in Section 5.

1.1 Related work

The literature survey focuses on the recently published work on an expert diagnostic system for hypertension risk stratification and detection of DM using finger PPG signal. We found that many researchers have proposed models that use morphological features. However, a combined study of an expert system that can perform a dual function of detecting hypertension and DM-II subjects is still little explored. Such a system can provide insight into the factors correlated with the overlap of HT and DM. The limitation recently reported an expert system that detects changes in the morphology of the PPG signal associated with the disease, but it requires a high-quality signal. To overcome the problem, researchers have reported a variety of filters such as Butterworth, Chebyshev

Elliptic filter, Wavelet de-noising filter, Chebyshev type I and type II, etc. (Liang et al., 2018b). Further, various types of signal quality indices were used to check signal quality, such as perfusion, skewness, kurtosis, entropy, signal-to-noise ratio, and relative power (Elgendi, 2016). The second issue in using morphological-based feature extraction is that it requires precise mapping of the fiducial point on the main PPG, which is obtained from zero crossings. In addition, the peak amplitude of the first and second derivative of PPG can pose a problem in the case of a noisy signal, so the derivative approach is unreliable in the case a noisy signal (Wei et al., 2018). The signal decomposition technique overcomes the limitation by decomposing signals into intrinsic mode functions (IMFs). Hadiyoso et al. (2020) used a Fourier transform to find the dominant spectral parameter associated with each IMFs.

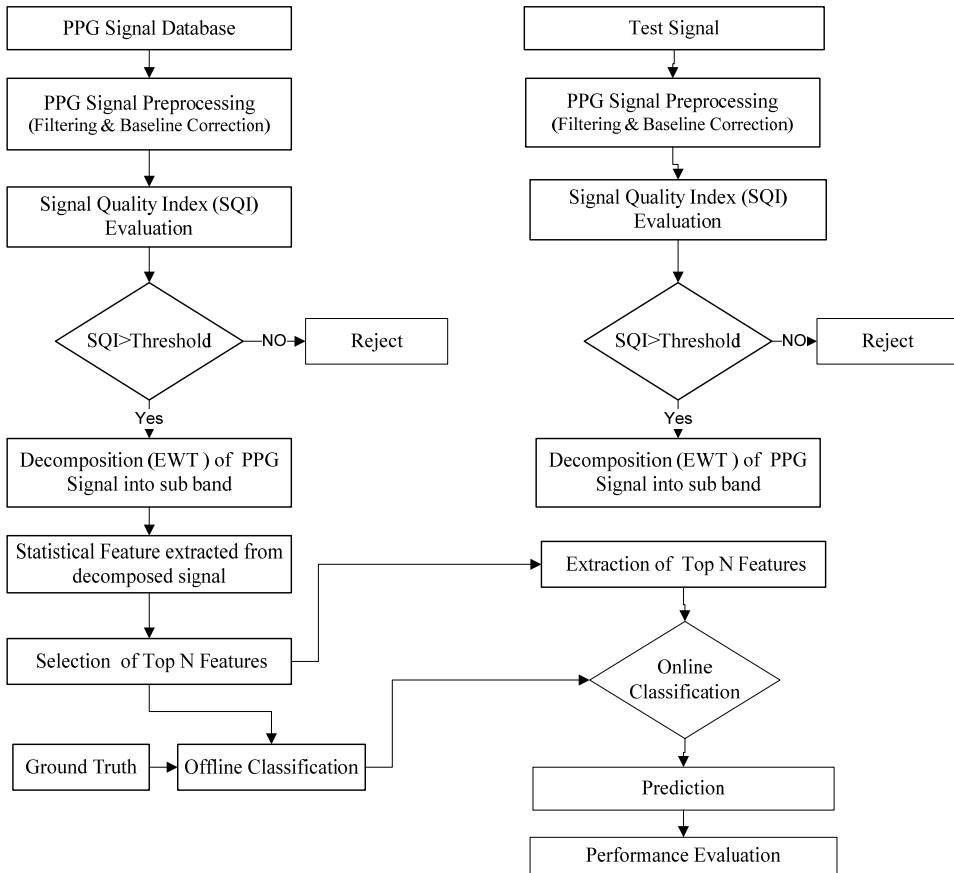
Carvalho et al. (2020) proposed a diagnostic model for seizure detection based on the decomposition of EEG signals into IMFs from which statistical features are extracted. The decomposition techniques, such as empirical mode decomposition (EMD), are used in a variety of diagnostic applications, such as studying arterial stiffness and vascular anomaly in diabetic patients (Wei et al., 2018). The advantage of the EWT decomposition method is that it involves the design of wavelet filter banks and decomposes a signal into a predetermined number of modes (Thirumala et al., 2020; Kumar et al., 2014) into compact bandwidth for each node in the Fourier spectrum, with central frequency. The vital advantage of the EWT is that it is less noise-sensitive. Recently soft computing has been widely used in various fields, such as healthcare systems, for its accuracy, reliability, and flexibility in design. One such model is proposed by Tjahjadi et al. (2020), which uses various soft computing techniques and features extracted using short-time Fourier transform (STFT) in the time domain to produce dominant features such as the instantaneous frequency and spectral entropy. The proposed model uses bidirectional long short-term memory (BLSTM) network to achieve higher accuracy as compared to other classifiers, i.e., convolutional neural networks (CNNs), KNN, bagged tree, logistic regression, and AdaBoost tree methods. The advantage of a hybrid feature selection algorithm is demonstrated by Zhang et al. (2005), in which a hybrid feature selection algorithm uses a GA combined with MI to select an optimal feature. The GA guide the search direction based on some criteria to find an optimal solution. The advantage of the approach is that it reduces dimensionality and improves performance. The primary motivation is to develop an expert diagnostic system for the detection of hypertension and DM-II detection. We conclude from the literature review that the EWT decomposition provides a robust model against noisy signals and a hybrid feature selection algorithm to select features that improve the model's accuracy.

2 Materials and methods

As shown in Figure 1, our proposed expert diagnostic system is used to detect hypertension and DM patients. The first block consists of a signal preprocessing module which performs operations such as PPG filtering, segmentation, and baseline correction. The second block evaluates the signal quality index (SQI) and selects the signal based on a predetermined threshold value of SQI. The selected, filtered signal is applied to the EWT decomposition block, which decomposes a signal into IMFs. Hilbert transform is applied to IMFs obtained from various decomposition levels that are from IMF0 to IMF8. A total of 99 statistical features are extracted from the resultant decomposed signal, a

total of 11 statistical features from each sub-band. A hybrid feature selection algorithm is used to select the optimal feature for hypertension as well as for DM-II detection. One of the two-hybrid feature selection methods is a majority voting method. The second hybrid algorithm is a combination of MI and GA, abbreviated as MI-GA, to select 16 features for categorical labels, normal prehypertension, hypertension type-1, and hypertension type-2. In contrast, only the majority voting method is used to choose a total of 16 features for a category normal versus DM-II. The selected feature is applied to the last block, which consists of a classification module. The classification module consists of a variety of classifiers such as KNN, GB, and sequential deep learning neural networks, which are optimised using the grid optimisation technique. Our proposed expert system is developed using Python programming language with the help of packages such as Keras, sklearn, and Scikit-learn (Abadi et al., 2015; Pedregosa et al., 2011).

Figure 1 Flowchart of the proposed expert diagnostic system for detection of HT and DM-II



2.1 Data collection

Dataset was acquired from DataCite Search figshare, an open-source clinical trial database. The database consists of three segments raw signal of 2.1 second PPG signal

recorded during 3 minutes, measurement cycle during which arterial systolic and diastolic BP value was measured using Omron HEM-7201 cuff based automatic digital blood pressure Monitor. The PPG sensor model used was dual light-emitting diodes (LED) with 660 nm and 905 nm wavelengths, with a sampling rate of 1 kHz, and the resolution ADC is 12-bit ADC. The dataset contains 657 PPG signal 2100 data points collected from 219 subjects, including normal, prehypertensive, hypertensive type I, and hypertensive type II W/o diabetes. In addition, the dataset contains information on demographic variables such as age, gender, Body Mass Index (BMI), and weight (Liang et al., 2018a).

2.2 Pre-processing

The raw PPG data filtered using the Chebyshev II bandpass filter in the 5th order with a cut-off frequency of 0.5 and 12 Hz is found to be optimal. The Chebyshev II filter provides excellent frequency selectivity, and the passband signal is less affected to retain the valuable information of the signal that is revealed in the decomposition of the signal into the sub-band. The performance of the filter is evaluated using Signal Quality Indices such as perfusion index (PSQI) and signal to noise (SNR) (Liang et al., 2018b).

3 Feature extraction

The EWT decomposes the PPG signal into its intrinsic modes, where each segmented signal is centred around resonance frequency within the boundaries of a frequency band. The EWT algorithm first determines the Fourier segments and then is followed by building a series of wavelet filters which are used to obtain different modes from the signal. The final phase operation determines the boundary of segmented, which contains the centre frequency known resonance frequency. To determine boundaries, the algorithm search for all the local maxima of the spectrum and then use the centre of the two adjacent local maxima as the boundaries of the Fourier segments signal. Finally, the Meyer wavelets filter is constructed as described by Gilles (2013), and a detailed description can be found in Literature (Thirumala et al., 2020). All other parameters were selected as per (Daubechies, 1992; Daubechies et al., 2011). EWT transforms the signal into IMFs into a set of narrow-band AM-FM components or IMFs, from which bandwidth is computed for amplitude modulation (BAM) and frequency modulation (BF M) component. Hilbert transform is applied on all IMFs using Welch window1024 to extract 11 features for each IMFs from IMF0 to IMF8, as mentioned below.

- 1 The bandwidth of amplitude modulation component: where A is the amplitude of the analytic signal, E is the energy and $\langle \omega \rangle$ is the centre frequency of the current mode, given by equation (1) (Cohen and Lee, 1990):

$$BW_{AM}^2 = \frac{1}{E} \int \left(\frac{dA}{dt} \right)^2 dt \quad (1)$$

- 2 The bandwidth of frequency modulation component:

$$BW_{FM}^2 = \frac{1}{E} \int \left(\frac{d\phi(t)}{dt} - \langle \omega \rangle \right)^2 A^2(t) dt \quad (2)$$

$\langle \omega \rangle$ is the centre frequency of the current mode, given by equation (12).

$$\langle \omega \rangle = \frac{1}{E} \int \frac{d\phi(t)^2}{dt} A^2(t) dt \quad (3)$$

- 3 Peak amplitude (Sp_{amp}): the peak amplitude of power spectral density (P_{XX}) for a particular bandwidth is given by

$$Sp_{amp} = \max(P_{XX}) \quad (4)$$

- 4 Peak frequency (Sp_{freq}) is the maximum frequency occurring in normalised P_{XX} is given by

$$Sp_{freq} = \arg \max(\overline{P_{XX}}) \quad (5)$$

- 5 Spectral entropy ($S_{Entropy}$): is calculated by applying the Shannon entropy concept to the power distribution of the Fourier-transformed signal to measure the uniformity of power spectral density underlying signal. The equation represents spectral entropy

$$S_{entropy} = - \sum_{f=0}^{\frac{f_s}{2}} \overline{P_{XX}}(f) \log[\overline{P_{XX}}(f)] \quad (6)$$

where $\overline{P_{XX}}$ denotes normalised spectral density (PSD) of the signal.

- 6 Spectral centroid (S_{Cent}): the spectral centroid is the centre of mass of the spectrum in which the spectrum distribution is defined as the weighted mean of the frequencies present in the signal. It is used to measure the spectral shape and position of the spectrum of the respective mode,

$$S_{Cent} = \frac{\sum_{f=0}^{\frac{f_s}{2}} \omega(f) M_{sp}(f)}{\sum_{f=0}^{\frac{f_s}{2}} M_{sp}(f)} \quad (7)$$

where f represent frequency bin, $\omega(f)$ and $M_{sp}(f)$ are central frequency and magnitude of the PSD of bin f (Huang et al., 2009).

- 7 Spectral energy: it measures signal energy distribution over the frequency band. It is given as:

$$\text{Spectral Energy } (S_{SE}) = \frac{1}{n} \sum_{f=0}^{\frac{f_s}{2}} P_{XX}(f) \quad (8)$$

where $P_{XX}(f)$ power spectral density is estimated by Welch's method, and N represents a total number of spectral coefficients (Welch, 1967; Gilles, 2013).

- 8 Skewness is related to the measured asymmetry or symmetry distributions of a signal, represented by the following equation:

$$S_{skew(x)} = E \left[\left(\frac{x(t) - \mu}{\sigma} \right)^3 \right] \quad (9)$$

where σ represents the standard deviation of signal $x(t)$, and it is mean denoted by μ .

- 9 Kurtosis is a measure related to the tail-heaviness of the distribution yielded by the signal and is given by

$$s_{kurt(x)} = E \left[\left(\frac{x(t) - \mu}{\sigma} \right)^4 \right] \quad (10)$$

- 10 Hjorth mobility: the Hjorth parameter was introduced in 1970 to study EEG signals in the time domain. The Hjorth parameter comprises a set of three parameters that are activity, mobility, and complexity. hjorth mobility is related to the mean frequency of the signal and is proportional to variance in the spectrum and expressed as a ratio per time unit. The Hjorth mobility is represented equation as given below (Vourkas et al., 2000; Navascués et al., 2009).

$$Hj_{Mob}(x) = \sqrt{\frac{Var\left(\frac{dx(t)}{dt}\right)}{var(x(t))}} \quad (11)$$

- 11 Hjorth complexity: a measure of the standard deviation of the slope with reference to the standard deviation of the amplitude.

$$Hj_{comp}(x) = \sqrt{\frac{Hj_{mob}\left(\frac{dx(t)}{dt}\right)}{Hj_{mob}(x(t))}} \quad (12)$$

The signal decomposition results in IMFs. The features extracted from IMFs are Differential entropy, Peak-magnitude of root-mean-square ratio (PRMS) (Kumar and Rao, 2019). Then these IMFs obtained are subjected to Hilbert transform to yield AM and FM frequencies. Then the second category of the features extracted are spectral peak magnitude peak frequency, spectral entropy and spectral energy (Marties et al., 2012):

3.1 Feature selection

The choice of feature selection technique can significantly impact the classification accuracy of the classifier. One of the feature selection techniques is called the filter method, which uses a ranking technique based on statistical measures. The limitation is that technique is classifier-independent. The second method, known as a wrapper method of feature selection, searches for a subset of features using a machine learning algorithm that employs a search strategy in which an optimal feature subset is selected with minimum size and is most informative to the dependent variables. But the fact is that most existing feature selection approaches which use greedy searching strategies and partial evaluation criteria can only produce local optimal feature subsets depending on the classifier used (Estévez et al., 2009).

We proposed a hybrid feature selection algorithm; the first algorithm uses a majority voting score. In this method, aggregation of five feature selection three feature selection methods from the filter method such as gain ratio (GR), symmetrical uncertainty (SU), and information gain (IG), while two from the wrapper method are consistency measure (C) and recursive feature elimination (RFE). Feature selected by any particular selection technique is assigned one vote. Finally, the feature with the highest number of votes assigned is selected (Singh et al., 2016; Chandrashekar et al., 2014). The second hybrid

method uses MI is used to evaluate fitness function (Brown, 2009) and GA used to search for an optimal feature. The summarised MI-GA algorithm is given below.

- Step 1 Create an initial population of individuals from the feature subset.
- Step 2 Compute the fitness of each individual and select an individual based on fitness.
- Step 3 Apply genetic operators to the population, such as reproduction, crossover, and mutation, and update the current population by replacing individuals based on fitness.
- Step 4 Loop step 2 UNTIL (stopping criteria).

3.2 Classification and evaluation methods

The classifier selection is a vital step in solving classification problems, data complexity, and some measures which can be investigated for the relation between the possible performance contribution and the complexity of the classifier so that optimal balance between performance, classifier complexity, and computational resources is required (Alceu et al., 2014). Our proposed expert diagnostic system uses three different classifiers: KNN, GB classifier, and sequential deep learning neural network.

Table 1 Features selected for hypertension and diabetes mellitus type-2

<i>S. no.</i>	<i>EWT features selected for hypertension</i>			<i>DM-II</i>	
	<i>Feature</i>	<i>Majority voting</i>	<i>Feature MI-GA</i>	<i>Feature</i>	<i>Majority voting</i>
1	AM3	5	pk3	Hcomb8	6
2	AM4	5	Hmob3	Hmob7	5
3	Cent3	4	freq4	freq2	5
4	ent4	4	Hcopmb8	Hcomp2	4
5	pk3	4	pow1	AM3	4
6	BM3	4	pow3	AM4	3
7	BM4	3	BM4	Cent3	3
8	Cent2	3	ent5	ent4	3
9	ent2	3	kurt5	pk3	3
10	ent3	3	Hmob7	BM3	3
11	freq2	3	pk8	BM4	3
12	Hcomp3	3	ent0	Cent2	3
13	Hmob3	3	Cent1	ent2	3
14	pk2	3	kurt1	ent3	2
15	pow3	3	BM2	freq2	2
16	skew2	3	freq2	Hcomp3	2

KNN is the nearest neighbours classifier, which classifies the feature to the nearest class based on its majority votes to its neighbours or because it stores all available observations and classifies each observation based on its similarity. KNN hyper-parameters is the number of nearest neighbours (K) included in a particular class. The weighting method

assigns weights to the data point, and distance metrics provide the distance between two elements of a class. GB is a machine learning technique that produces a prediction model in the form of an ensemble of weak prediction models. In the case of GB, the new models fit to deliver a more accurate estimate of the class variable by the learning method. Every new model is correlated with the negative gradient of the loss function of the system and tends to minimise it by using the gradient descent method (Bynagari and Ahmed, 2021). GB classifier is widely used for its efficiency, accuracy, and by training, the errors associated with a slight gradient to discard every instance related to small gradients (Rapp, 2021). Sequential Deep Learning neural networks provide flexibility in design. The input consists of a 1D sequence, two or more hidden dense layers, and an output layer. The Rectilinear function (Relu), sigmoid, or softmax function can be used in hidden and output layers. The function of dropout layer applies a mask to the inputs for disconnecting some of the neurons. The advantage of using the dropout layer is that it avoids the fitting problem (Nowak et al., 2020; Wang et al., 2018). The classifiers are trained and tested using 10-fold cross-validation for performance evaluation. The comparison of hypertension risk stratification models is conducted using F1-score parameter and one vs. one classification approach.

4 Results and discussion

The result obtained for the classification of hypertension for category NT versus PHT, NT versus HT-I, NT vs. HT-II, two feature selection techniques such as majority voting score and MI-GA are shown in Table 2. The maximum accuracy achieved in terms of F1 score for category NT versus PHT, NT versus HT-I, and NT vs. HT-II is 78.9%, 91.2%, and 94%, using soft computing based classifier and hybrid feature selection technique.

The maximum accuracy achieved for category NT versus HT-II in terms of F1 score is 94% using GB classifier with a learning rate (lr_rate) of 0.01, maximum depth of split of 3, and the optimal number of estimators (n_estimators) is found to 100 which requires significant computational time. The hybrid feature selection technique based on the majority voting score feature selection algorithm improves the performance compared to the MI-GA. The features selected using two feature selection techniques mainly contains statistical feature extracted from IMF2 to IMF4. Hilbert transform (HT) is a form of integral transformation that provides a direct examination of vibrations in terms of instantaneous attributes: frequency, phase, and amplitude to be analysed in the time domain and the frequency domain (Feldman, 2001). The grid search algorithm is simple to implement and efficient when the number of features is small in number (Chandrashekar et al., 2014). The sequential deep neural network designed for classification uses 1D data and six dense layers with rectified linear unit (Relu) activation function. The first four layers are dense, and the last two have a sigmoid activation function. The number of neurons is 30 in the input, 80 in the subsequent two layers, and the number of neurons in the last three layers is 40, 8, and 1, respectively, at the output layer (Siddiqi, 2019). Adam optimisation algorithm is for sequential deep learning neural network which adapts the learning rate to different parameters automatically, based on the statistics of a gradient. It simplifies learning rate settings, leading to faster convergence (Zhang, 2018). The number of epochs is kept to 150 for training and loss function based on cross-entropy. During tuning, the GB classifier cross-validation is used for early stopping and a low learning rate to reduce the chance of overfitting. At the same

time, subsampling of less than 100% of the training observations minimises the chances of overfitting. From the observation in Table 2, the classification accuracy of NT versus prehypertension is found to be lowest compared to category NT vs. HT type I and NT vs. HT type II. Our proposed model has shown robustness to low-quality signals.

Table 2 Result of the HT classification using the majority voting score method and MI-GA feature selection technique

<i>Classifier/kernel parameters/optimisation technique</i>	<i>TP</i>	<i>FN</i>	<i>FP</i>	<i>TN</i>	<i>SE%</i>	<i>SP%</i>	<i>ACC%</i>	<i>AUC</i>	<i>Precision</i>	<i>F1%</i>
<i>Normal vs. prehypertension (feature selection: majority voting method) Normal sample=240 hypertension type I sample = 105 CV = 10 fold (classification strategy: one vs one)</i>										
KNN metric': 'Euclidean', neighbours: 19, weights: uniform	134	106	65	193	55.8	74.8	66	0.7	0.657	61
GradientBoostingClassifier, lr_rate: 0.01, max_depth 3, n_estimators: 50, subsample: 0.7	125	115	43	215	52.1	83.3	68	0.78	657	61
Sequential deep learning NN	173	67	37	221	72.1	85.7	79	0.83	0.795	78.9
<i>Normal vs. prehypertension (feature selection:-MI-GA)</i>										
KNN, ('Euclidean', 'n_neighbours': 19, 'weights': 'uniform)	148	92	73	185	66.97	61.67	67	0.72	0.669	64.2
GradientBoostingClassifier lr: 0.01, max_depth': 3, estimators': 50, 'subsample': 0.5	120	120	40	218	75.	50.	68	0.78	0.679	60
Sequential deep learning NN	184	56	52	206	77.97	76.67	78	0.81	0.783	78.9
<i>Normal vs. hypertension type I (feature selection:-majority voting method) normal sample = 240 hypertension type I sample = 105 CV=10 fold (classification strategy :one vs one)</i>										
KNN metric':manhattan', neighbours': 5, 'weights': 'uniform'	216	24	48	57	90.0	54.2	79	0.86	0.79	85.7
Gradient BoostingClassifier, lr_rate: 0.01, max_depth 3, n_estimators:50 subsample:0.7	237	3	43	62	98.7	59.0	87	0.97	0.84	91.2
Sequential deep learning NN	215	25	12	93	89.5	0.89	88	0.97	0.86	87.7

Table 2 Result of the HT classification using the majority voting score method and MI-GA feature selection technique (continued)

<i>Classifier/kernel parameters/optimisation technique</i>	<i>TP</i>	<i>FN</i>	<i>FP</i>	<i>TN</i>	<i>SE%</i>	<i>SP%</i>	<i>ACC%</i>	<i>AUC</i>	<i>Precision</i>	<i>F1%</i>
<i>Normal vs. hypertension type I (feature selection: MI-GA)</i>										
KNN (metric': 'Mannhattan', 'n_neighbours': 5, 'weights': 'uniform)	219	21	44	61	83.2	91.2	81	0.87	0.74	65.2
Gradient BoostingClassifier lr: 0.1, max_depth': 3, estimators': 10, 'subsample': 1.0	237	7	60	45	79.8	97.1	82	0.91	0.81	88
Sequential deep learning neural network	226	14	23	82	90.7	94.1	89	0.91	0.88	88
<i>Normal vs. hypertension type II (feature selection:-majority voting method) normal sample =240 hypertension type II sample=54 CV=10 fold (classification strategy :one vs one)</i>										
KNN metric': 'Euclidean', neighbours: 5, weights: uniform	233	7	36	18	86.6	97.0	85	0.89	0.85	92
GradientBoostingClassifier, lr_rate: 0.01, max_depth 3, n_estimators:100 subsample:0.5	239	1	32	22	99.5	37.0	89	0.97	0.89	94
Sequential deep learning neural network	234	6	17	37	97.5	68.5	92	0.92	0.89	85.8
<i>Normal vs. hypertension type II (feature selection: MI-GA)</i>										
KNNmetric': 'Euclidean', 'n_neighbors':7, 'weights': uniform)	236	4	32	22	88.0	98.33	88	0.93	0.88	92.9
GradientBoostingClassifier lr: 0.01, max_depth': 3, estimators': 100, 'subsample': 0.7	240	0	46	8	83.9	100	84	0.9	0.84	91.3
Sequential deep learning NN	228	2	36	18	86.3	99.1	87	0.89	88	70

Table 3 Result of classification normal vs. DM-II

<i>Classifier/kernel parameters/optimisation technique</i>	<i>TP</i>	<i>FN</i>	<i>FP</i>	<i>TN</i>	<i>SE%</i>	<i>SP%</i>	<i>ACC%</i>	<i>AUC</i>	<i>Precision</i>	<i>F1%</i>
GradientBoostingClassifier, {'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 50, 'subsample': 0.5}	355	50	11	64	91	87.3	87	78	0.96	92
Sequential deep learning neural network	398	7	10	65	98.51	8.33	96	0.97	0.97	97.9

4.1 Result of classification of diabetes mellitus patients

Table 3 shows the results of the classification of normal vs. DM-II. A sequential deep learning neural network achieves the maximum F1 score of 97.9%, as depicted in Table 3. The majority voting score features selection is used to select 16 top score features, as depicted in Table 1. It is observed that Hcomp 8 and Hmob 7 features are selected, which represent statistical features from a higher frequency band. Therefore, the classifier's performance is not significantly affected by class imbalance.

5 Conclusions

The main advantage of our proposed expert diagnostic model is that it offers dual functions for detecting hypertensive patients and subjects suffering from DM, which can be helpful in the clinical environment for large-scale screening and reduced diagnostic time. Ours proposed a framework that combines the advantages of EWT and Hilbert transform to extract statistical features from PPG signal, then uses hybrid feature selection algorithms to find feature subsets, improving the classification accuracy. EWT transforms the signal into narrowband IMF, which makes the model robust to noisy signals. Sequential deep learning neural networks provide flexibility in designing networks from simple to complex to solve a complex classification problem. The main contribution of the proposed model is to provide high accuracy for hypertension risk classification and accurately detect DM-II patients simultaneously. However, the limitation of our work dataset was not large enough, and there was a class imbalance. We plan to extend our work to improve the proposed methodology in terms of an efficient design of deep learning networks to improve the classification of the model on a larger dataset for strong generalisation.

References

- Abadi, M. et al. (2015) *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, DOI: 10.5281/zenodo.4724125.
- Akhtar, I. et al. (2018) 'Higher-order spectral analysis to identify quadratic nonlinearities in fluid-structure interaction', *Journal Mathematical Problems in Engineering*, Hindawi, DOI: 10.1155/2018/2394124.
- Alceu, S. et al. (2014) 'Ensemble of classifiers, dynamic selection of classifiers, data complexity', *Journal of Pattern Recognition*, Vol. 47, DOI: <https://doi.org/10.1016/j.patcog.2014.05.003>.
- Brown, G. (2009) 'A new perspective for information theoretic feature selection', *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics (AISTATS)*, Clearwater Beach, Florida, USA, Vol. 5, JMLR: W&CP 5 [online] <https://proceedings.mlr.press/v5/brown09a.html>.
- Bundy, J.D et al. (2019) 'Comparison of the 2017 ACC/AHA hypertension guideline with earlier guidelines on estimated reductions in cardiovascular disease', *Current Hypertension Reports*, Vol. 21, No. 10, p.76, <https://doi.org/10.1007/s11906-019-0980-5>.
- Bynagari, N.B. and Ahmed, A.A.A. (2021) 'Anti-money laundering recognition through the gradient boosting classifier', *Academy of Accounting and Financial Studies Journal*, Vol. 25, No. 5, Print ISSN: 1096-3685; Online ISSN: 1528-2635.
- Carvalho, V.R. et al. (2020) 'Biomedical signal processing and control evaluating five different adaptive decomposition methods for EEG signal seizure detection and classification', *Biomedical Signal Processing and Control*, Vol. 62, DOI: 10.1016/j.bspc.2020.102073.

- Castaneda, D. et al. (2018) 'A review on wearable photoplethysmography sensors and their potential future applications in health care', *Int J Biosens Bioelectron*, Vol. 4, No. 4, pp.195–202, DOI: 10.15406/ijbsbe.2018.04.00125. PMID: 30906922; PMCID: PMC6426305.
- Chandrashekar, G. et al. (2014) 'A survey on feature selection methods', *Computers & Electrical Engineering*, Vol. 40, No. 1, pp.16–28, ISSN 0045-7906, <https://doi.org/10.1016/j.compeleceng.11.024>.
- Cohen, L. and Lee, C. (1990) 'Instantaneous bandwidth for signals and spectrogram', *Proc. IEEE Int. Conf. Acoustic, Speech, Signal Process*, pp.2450–2454, DOI: 10.1109/ICASSP.1990.116086Corpus ID: 62219338.
- Daubechies, I. (1992) 'Ten lectures on wavelets', *CBMS-NSF Regional Conference Series in Applied Mathematics*, Philadelphia, PA, USA, SIAM, <https://doi.org/10.1137/1.9781611970104>.
- Daubechies, I. et al. (2011) 'Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool', *Applied and Computational Harmonic Analysis*, Vol. 30, No. 2, pp.243–261, ISSN 1063-5203, <https://doi.org/10.1016/j.acha.2010.08.002>.
- Elgendi, M. (2016) 'Optimal signal quality index for photoplethysmogram signals', *Bioengineering (Basel, Switzerland)*, Vol. 3, No. 4, p.21, <https://doi.org/10.3390/bioengineering3040021>.
- Elgendi, M. et al. (2019) 'The use of photoplethysmography for assessing hypertension', *NPJ. Digit. Med.*, Vol. 2, p.60, <https://doi.org/10.1038/s41746-019-0136-7>.
- Estévez, P.A. et al. (2009) 'Normalized mutual information feature selection', *IEEE Trans. on Neural Network*, 2 February, Vol. 20, pp.189–201, Accession Number: 10471900, DOI: 10.1109/TNN.2008.2005601.
- Feldman, M. (2001) 'Hilbert transforms', in *Encyclopedia of Vibration*, pp.—48, Elsevier, <https://doi.org/10.1006/rwvb.2001.0057>.
- Gilles, J. (2013) 'Empirical wavelet transform', *IEEE Transactions on Signal Processing*, Vol. 61, pp.3999–4010, <http://dx.doi.org/10.1109/TSP.2013.2265222>.
- Hadiyoso, S. et al. (2020) 'Comparison of EMD, VMD and EEMD methods in respiration wave extraction based on PPG waves', *Journal of Physics: Conference Series*, DOI:10.1088/1742-6596/1577/1/012040.
- Huang, C.J., Yang, Y.J., Yang, D.X. and Chen, Y.J. (2009) 'Frog classification using machine learning techniques', *Exp. Sys. with App.*, Vol. 36, No. 2, Part 2, pp.3737–3743.
- Huang, N.E. et al. (1998) 'The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis', *Proc. R. Soc. Lond. A*, 454903–995, <https://doi.org/10.1098/rspa.1998.0193>.
- Keerthika, T. et al. (2019) 'An effective feature selection for heart disease prediction with aid of hybrid kernel SVM', *International Journal of Business Intelligence and Data Mining*, Vol. 15, No. 3, pp.306–326, DOI: 10.1504/IJBIDM.2019.101977.
- Kumar, M.R. and Rao, Y.S. (2019) 'Epileptic seizures classification in EEG signal based on semantic features and variational mode decomposition', *Cluster Comput.*, Vol. 22, pp.13521–13531, <https://doi.org/10.1007/s10586-018-1995-4>, DOI: 10.1109/TNN.2008.2005601
- Kumar, R. et al. (2014) 'Empirical wavelet transform based ECG signal compression', *IETE J. Res.*, Taylor & Francis, Vol. 60, pp.423–431, DOI:10.1080/03772063.2014.963173.
- Liang, Y. et al. (2018a) 'A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China', *Sci Data*, Vol. 27, No. 5, 180020, DOI: 10.1038/sdata.2018.20. PMID: 29485624; PMCID: PMC5827692.
- Liang, Y. et al. (2018b) 'Analysis: an optimal filter for short photoplethysmogram signals', *Scientific Data*, Vol. 5, pp.1–12, DOI: 10.1038/sdata.2018.76.
- Martis, R.J. et al. (2012) 'Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals', *Int J Neural Syst.*, December, Vol. 22, No. 6, 1250027, DOI: 10.1142/S012906571250027X.

- Navascués, M.A. et al. (2009) 'Time domain indices and discrete power spectrum in electroencephalographic processing', *Int. J. Comput. Math.*, Vol. 86, pp.1968–1978, DOI:10.1080/00207160903023599.
- Nowak, J. et al. (2020) 'Discovering sequential patterns by neural networks' *International Joint Conference on Neural Networks (IJCNN)*, pp.1–6, DOI: 10.1109/IJCNN48605.2020.9207461.
- Pedregosa, F. et al. (2011) 'Scikit-learn: machine learning in Python', *Journal of Machine Learning Research*, 1 November, Vol. 12, pp.2825–2830, JMLR.org, ISSN:1532-4435.
- Petrie, J.R. et al. (2017) 'Diabetes, hypertension, and cardiovascular disease: clinical insights and vascular mechanisms', *Can J. Cardiol.*, May, Vol. 34, No. 5, pp.575–584, DOI: 10.1016/j.cjca.2017.12.005. Epub 2017 December 11, PMID: 29459239; PMCID: PMC5953551.
- Rapp, M. (2021) 'Machine learning, multi-label classification, gradient boosting, rule learning', *Software Impacts*, Vol. 10, Elsevier, <https://doi.org/10.1016/j.simpa.2021.100137>.
- Shaker, B.H. et al. (2019) 'Grid search-based hyper parameter tuning and classification of microarray cancer data', *Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*, pp.1–8, DOI: 10.1109/ICACCP.2019.8882943.
- Siddiqi, R. (2019) 'Automated pneumonia diagnosis using a customized sequential convolutional neural network', *Proceedings of the 2019 3rd International Conference on Deep Learning Technologies (ICDLT 2019)*, Association for Computing Machinery, New York, NY, USA, pp.64–70, DOI: <https://doi.org/10.1145/3342999.3343001>.
- Singh, B.K. et al. (2016) 'Fuzzy cluster based neural network classifier for classifying breast tumors in ultrasound images', *Expert Syst. Appl.*, Vol. 66, pp.114–23, <http://dx.doi.org/10.1016/j.eswa.2016.09.006>.
- Thirumala, K> et al. (2020), Empirical wavelet transform and dual feed-forward neural network for classification of power quality disturbances, *Int. J. Power and Energy Conversion*, Vol. 11, No. 1, DOI: 10.1504/IJPEC.2020.104805.
- Tjahjadi, H. et al. (2020) 'Non-invasive classification of blood pressure based on photoplethysmography signals using bidirectional long short-term memory and time-frequency analysis', *IEEE Engineering in Medicine and Biology Society Section*, DOI: 10.1109/ACCESS.2020.2968967.
- Vourkas, M. et al. (2000) 'Use of ANN and Hjorth parameters in mental-task discrimination', *First International Conference on Advances in Medical Signal and Information Processing*, IEE Conf Publ No. 476, pp.327–332, 10.1049/cp:20000356.
- Wang, L. et al. (2014) 'Feature selection based on meta-heuristics for biomedicine', *Optimization Methods and Software*, Vol. 29, pp.703–719. DOI:10.1080/10556788.2013.834900.
- Wang, Y. et al. (2018) 'A survey and formal analyses on sequence learning methodologies and deep neural networks', *2018 IEEE 17th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)*, pp.6–15, DOI: 10.1109/ICCI-CC.2018.8482022.
- Wei, H-C. et al. (2018) 'Instantaneous frequency from Hilbert-Huang transformation of digital volume pulse as an indicator of diabetes and arterial stiffness in upper-middle-aged subjects', *Sci. Rep.*, Vol. 8, 15771, <https://doi.org/10.1038/s41598-018-34091-6>.
- Welch, P. (1967) 'The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms', *IEEE Trans Audio Electroacoust.*, Vol. 15, pp.70–73, DOI:10.1109/TAU.1967.1161901.
- Zhang, Z. (2018) 'Improved adam optimizer for deep neural networks', *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, pp.1–2, DOI: 10.1109/IWQoS.2018.8624183.
- Zhang, C.K. et al. (2005) 'An effective feature selection scheme via genetic algorithm using mutual information', in Wang, L. and Jin, Y. (Eds.): *Fuzzy Systems and Knowledge Discovery. FSKD 2005. Lecture Notes in Computer Science*, Vol. 3614, Springer, Berlin, Heidelberg, https://doi.org/10.1007/11540007_10.