Automated Detection of Hypertension Disease Using Machine Learning and Signal Processing-Based Methods



Jaypal Singh Rajput and Manish Sharma

Abstract Hypertension is a critical health issue and an important area of research because of its high pervasiveness and a remarkable risk factor for cardiovascular and cerebrovascular disease. However, it is a silent killer in all respects. Hardly any side effect can be found in its beginning period until an extreme Medical emergency like heart attack, stroke, or chronic kidney disease. Since individuals are unaware of hypertension, the identification is possible through measurement only. The detection of hypertension at the beginning stage can protect from serious health issues. Furthermore, the hypertension diagnosis by measuring blood pressure may not reflect any severe complication caused due to high blood pressure. Alternatively, automated machine learning and signal processing-based methods require to detect hypertension and its complication (syncope, stroke, and myocardial infraction) from the direct ECG signal. However, the ECG signal is non-stationary, and experts may commit mistakes in observation. As a result, the delay in the treatment of hypertension can be life threatening. Therefore, we have developed the automated detection algorithm for HPT-influenced electrocardiogram (ECG) signal using an optimal filter bank and machine learning. A total of six sub-bands were produced from each ECG signal using a filter bank. In addition, we have extracted the various linear and nonlinear features for all six sub-bands. Subsequently, a ten-fold cross-validation technique was employed for the k-nearest neighbor (KNN) classifier to classify the ECG signals. As a result, the proposed model has achieved a classification accuracy of 98.4%. Hence, the proposed work classifies hypertension from ECG signals in myocardial infarction, stroke, syncope, and low-risk hypertension. Moreover, we can install the proposed algorithm on a personal computer and diagnose the HPT-associated disease from an ECG signal.

Keywords Hypertension · ECG signal · Machine learning · Classification · Filter bank · Wavelet decomposition · Signal processing

J. S. Rajput (🖂) · M. Sharma

Department of Electrical and Computer Science Engineering, Institute of Infrastructure Technology Research and Management, Ahmedabad 380026, India e-mail: jaypal2020@gmail.com

[©] The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022 R. N. Shaw et al. (eds.), *Advanced Computing and Intelligent Technologies*, Lecture Notes in Electrical Engineering 914, https://doi.org/10.1007/978-981-19-2980-9_4

1 Introduction

Hypertension (HPT) is a critical health issue. It can severely affect human health, and its pervasiveness increases on a global level, but the rate of HPT awareness, treatment, and control remains slow [11]. The World Health Organization (WHO) is more vigilant of HPT treatment, understanding, and diagnosis [11]. HPT is defined when the systolic and diastolic blood pressure is greater than 140/80 in more than three clinical trials.

Hypertension is a remarkable state that can indicate many severe diseases like stroke (STR), syncope (SYN), myocardial infarction (MI), and heart disease [6]. Blood pressure (BP), smoking, overweight, lack of exercise, excessive salty eating, stress, age, family ancestry, kidney illness, and thyroid disease are the few reasons that cause hypertension [24].

The functionality of the heart is recorded by an ECG signal in the form of an electrical signal. Therefore, the ECG signal is more relevant in HPT detection and the disease associated with it [10]. Simjanoska et al. [22] specify the relationship between ECG and blood pressure and how ECG changes when the blood pressure is changed.

The primary motivation of this research work is to detect hypertension-associated diseases from the ECG signal. However, early detection of hypertension can save many lives and enhance people's life quality.

Various devices, methods, and algorithms have been developed to detect hypertension. Similarly, the details of work done on HPT detection in literature are mentioned below:

Rajput et al. [8] discriminate the severity of hypertension ECG signal using hypertension diagnosis index (HDI). The developed HDI model classifies the low and high-risk hypertension ECG signals with 100% classification accuracy.

In another study [10], the authors classify the low, high-risk hypertension, and normal ECG signals using signal processing and machine learning-based methods. In addition to this, they obtained 99.95% classification accuracy using the ensemble bagged tree classifier.

Further, in the subsequent study [18], the classification accuracy of 98.05% was obtained without wavelet-based methods. The classification has been conducted on the severity of hypertension and normal ECG signals.

Moreover, in another study [9], they have classified hypertension and normal Balistocardiogram signals using empirical mode decomposition and wavelet transform methods. As a result, the authors obtained the highest classification accuracy of 87%.

Quachtran et al. [7] extract intracranial pressure (ICP) from ECG signal and developed a deep learning model for the detection of intracranial hypertension. The deep learning model gives $92.0 \pm 2.25\%$ accuracy.

Sau et al. [12] worked on seafarer people's depression and anxiety using machine learning. The precision and accuracy of their developed model are 82.6% and 84%, respectively.

In another study, Melillo et al. [5] designed an automated detection of high-risk hypertension algorithm from heart rate variability (HRV) signals using machine learning methods. The sensitivity and specificity of HRV-based models are 71.4 % and 87.8 %, respectively.

Ni et al. [6] employ a HRV signal-based multi-scale fine-grained model to detect the severity of hypertension. The HRV signal-based model gives 95% accuracy using a machine learning algorithm.

Simjanoska et al. [22] identified the SBP and DBP from ECG signal using machine learning methods. The SBP and DBP achieved 9.45 and 8.13 mmHg mean absolute error.

Song et al. [23] distinguish hypertension and heart disease from the HRV signal. In addition, the Naive Bayes classifier, a machine learning-based model, gives classification accuracy of 92.3%.

Hence, it is apparent from the literature, and disease (STR, SYN, MI, and LHT) associated with hypertension has not been studied yet. Therefore, in the current scenario, hypertension is attracting researchers globally. To diagnose and predict hypertension and its associated disease, a large amount of recorded data is available in hospitals and online databases. Accordingly, we have developed the hypertension diseases detection system by signal processing and machine learning-based methods. In the proposed work, we have used an orthogonal wavelet filter bank (OGWFB) to perfectly discriminate STR, SYN, MI, and LHT ECG signals. The OGWFB produces six sub-bands (SBs) for each ECG signal considering five-level wavelet decomposition. In addition, the LOGE and SLFD features were calculated for all SBs. As a result, the KNN classifier presents the highest classification accuracy of 98.4%.

Section 2 provides the details of the dataset. Then, the methodology is explained in Sect. 3. Subsequently, the performance (result) of the developed model is discussed in Sect. 4. At last, the outcome and concluding remarks of the proposed algorithm are given in Sect. 5.

2 Dataset

The dataset for this research work was obtained from Physionet's online database (SHAREE database). The Ethics Committee approved the current study of Federico II University Hospital Trust. A total of 139 hypertensive recordings were used, out of which 49 are female, and 90 are male patients; the average age is 55 years. In addition to this, the length of each ECG signal is 2h:10 min:12-s. Furthermore, each ECG signal has III, V3, V5 leads and approximately one million samples (samples/signal). The lead III, V3, and V5 are assigned as CH1 (channel1), CH2 (channel2), and CH3 (channel3). The ECG signal sampling frequency, bit resolution, and sampling intervals are 128 Hz, 8-bit, and 0.0078125 sec, respectively. Out of 139 subjects, three are SYN, three are STR, and 11 are MI, while 122 patients are low-risk hypertension LHT subjects. Further, we have segmented each ECG signal into a 5-minute duration signal. After segmentation of ECG signal, 3172 ECG signals are of LHT, 78 of stroke,



Fig. 1 Low-risk hypertension ECG signal



Fig. 2 Myocardial infarction ECG signal

78 of syncope, and 286 of myocardial infarction. Figures 1, 2, 3, and 4 show all four classes of hypertension-associated ECG signals of 5-min duration.

3 Methodology

The optimally designed OGWFB discriminates LHT, MI, STR, and SYN classes of ECG signal. Each ECG signal has been decomposed into various sub-bands using a filter bank. The LOGE and SLFD features were computed for each ECG signal SBs. As a result, a total of 12 (six LOGE and six SLFD) features were obtained from each ECG signal. Subsequently, we applied various machine learning algorithms



Fig. 4 Syncope ECG signal

on features calculated ECG signals. The KNN machine learning classifier gives the highest accuracy. The outline of the developed algorithm is presented in Fig. 5.

3.1 Preprocessing of ECG Signal

Z-score normalization is performed on each epoch of the ECG signals to eliminate the amplitude scaling problem [1, 3, 4]. Five-minute ECG signals are generated by segmenting the long-length ECG signal.

3.2 OGWFB Filter Bank

The two-band filter bank has an analysis filter bank (decomposition) and synthesis filter (reconstruction) is shown in Fig.6. Analysis filter bank has $P_0(z)$ low-pass



filter and $P_1(z)$ high-pass filter. The high- and low-pass analysis filter bank output is down-sampled by a factor of 2, while synthesis filter bank input is up-sampled by a factor of 2. In the synthesis filter bank, $Q_0(z)$ is low pass and $Q_1(z)$ is high-pass filter. In the proposed work, we used an orthogonal wavelet filter bank developed by [15]. The output of the synthesis filter bank is matched with the input to the analysis filter bank to get the same result.

Perfect reconstruction is achieved using the two-channel filter bank. However, the condition of orthogonality is necessary for filters to get perfect reconstruction of signal [2, 16, 26, 28, 30, 31]. Therefore, the orthogonal filter bank can be con-

Fig. 5 Layout of the proposed work



Fig. 6 Orthogonal wavelet filter bank diagram

verted into the finite impulse response analysis low-pass filter $p_0(n)$, and it must fulfill the condition of orthogonality which is equivalent to the condition of perfect reconstruction and zero moments [21, 29]. Additionally, the high-pass filter $p_1(n)$ can be produced by adjusting the sign of the coefficient of the flipped variant of the low-pass filter. Finally, the synthesis bank filters can be extracted from the time reversal of the analysis banks.

3.3 Wavelet Decomposition

The ECG signal is non-stationary; therefore, we cannot apply conventional (Fourier, Laplace, and short-time Fourier transform) methods [16, 19, 21]. Instead, we used an optimal wavelet filter bank (OGWFB) to decompose ECG signals in various subbands. In addition, a five-level wavelet decomposition was used [14, 27]- [21, 34]. As a result, it produces accurate and precise information about the ECG signals. Total N + 1 sub-bands were made for N level wavelet decomposition [14, 17]. However, the SB1-SB5 are detailed, and SB6 is an approximate sub-band.

3.4 Features Used in Proposed Work

The important part of this work is to calculate and select the required features. Significantly, the performance of the classifier is based on the nature of the feature extracted. Moreover, it is not priory known which feature will best discriminate each class of ECG signal. In addition, the LOGE and SLFD features were computed for all six SBs of each ECG signal [33]. Finally, the feature extracted ECG signals were fed to the machine learning classifiers. As a result, we can classify the LHT, MI, SYN, and STR ECG signals using LOGE and SLFD features with machine learning classifiers.

3.5 Classification and Performance Evaluation

In the proposed work, we have used various supervised machine learning algorithms for the automated classification of hypertension ECG signals. However, we use ECG signals on a variety of classifiers, including support vector machines (SVM), k-nearest neighbor (KNN), decision trees (DT), and ensemble bagged trees (EBT), to improve performance. As a result, we have obtained the highest classification performance using the KNN classifier.

Table 1 Terrormance summary of mens was obtained using errit, erriz, and errs						
Filters	CH1, ACC (%)	CH2, ACC (%)	CH3, ACC (%)			
<i>F</i> 1	98	97.4	98.4			
F2	98.2	97.1	98.4			
F3	98.2	97.1	98.3			

 Table 1
 Performance summary of filters was obtained using CH1, CH2, and CH3

Table 2 Comparison of AUC obtained by all fill	litters for CH1.	, CH2 and C	ιнэ
---	------------------	-------------	-----

Filters	CH1, AUC	CH2, AUC	CH3, AUC
<i>F</i> 1	0.95	0.96	0.98
F2	0.96	0.97	0.99
F3	0.96	0.95	0.98

Table 3	Confusion	matrix of	CH1	for all	classes	obtained	using	KNN	classifier
---------	-----------	-----------	-----	---------	---------	----------	-------	-----	------------

True class	Predicted class	Predicted class					
	Normal, 0	STR, 1	SYN, 2	MI, 3			
Normal, 0	3143	6	4	19			
STR, 1	7	71	0	0			
SYN, 2	3	0	75	0			
MI, 3	24	2	1	259			

True class	Predicted class					
	Normal, 0	STR, 1	SYN, 2	MI, 3		
Normal, 0	3131	9	14	18		
STR, 1	7	70	0	1		
SYN, 2	21	0	57	0		
MI, 3	24	0	0	262		

Table 4 Confusion matrix of CH2 for all classes obtained using KNN classifier

Table 5 Confusion matrix of CH3 for all classes obtained using KNN classifier

True class	Predicted class				
	Normal, 0	STR, 1	SYN, 2	MI, 3	
Normal, 0	3149	1	5	17	
STR, 1	3	74	0	1	
SYN, 2	8	0	70	0	
MI, 3	23	1	0	262	

Table 6 Classification accuracy of various classifiers for CH1, CH2, and CH3

Classifier	CH1, ACC (%)	CH2, ACC (%)	CH3, ACC (%)
KNN	98.2	97.4	98.4
DT	91.0	90.1	92.4
EBT	93.5	92.6	93.4
SVM	97.3	96.0	96.7

Usually, KNN is applied for dimensionality reduction and feature selection [13, 20, 25]. In addition, the KNN is used for the k training samples, which are neighbors of the test sample, to classify it. Following this, the KNN classifier provided the lowest probability and overfitting [13, 20, 25].

4 Result and Discussion

The experimental work has been performed on the MATLAB version (9.1.0), with Intel Xeon 3.5 GHz, and 16 GB RAM. The F1, F2, and F3 filters of OGWFB enhance the performance of the proposed work. In addition, filter F2 presents the highest classification accuracy compared to the other two filters. The performance of each filter in terms of classification accuracy is shown in Table 1. A filter F2 produced the highest area under the curve (AUC) of 0.99 for CH3 is mentioned in Table 2. Tables 3, 4, and 5 represent the confusion matrix of LHT, MI, STR, and SYN of CH1, CH2, and CH3 for KNN classifiers. The classification performance of each classifiers is



presented in the Table 6. It is evident from Table 6 that the KNN classifier givesclassification accuracy of 98.4% for all classes. For testing the model performance and avoiding overfitting, we used the ten-fold cross-validation method. Table 6 shows that our model can identify 98.4% accurately of LHT, MI, STR, and SYN classes. The best receiver operating characteristics (ROC) curve and AUC of KNN classifier are shown in Figs. 7, 8, and 9 for CH1, CH2, and CH3.

5 Conclusion

This study used optimal OGWFB to separate LHT, MI, STR, and SYN ECG signals. The LOGE and SLFD features were calculated for all six sub-bands. OGWFB



can classify ECG signals accurately using LOGE and SLFD with a ten-fold crossvalidation method. To check the performance of filter banks, we have applied various classifiers. KNN classifier presents an accuracy of 98.4% and an AUC of 0.99. This study can be employed for the identification of heart, brain, and kidney disease. Therefore, an adaptable, robust, reliable, and accurate model has been proposed. Furthermore, the system performance can be enhanced by extracting other features like signal sample entropy, wavelet entropy, and higher-order spectra. These findings demonstrate that our methods outperform previous models and that they can be used in large databases. Sequentially, we can test performances of the suggested technique on a large dataset for automatic detection of the severity of hypertension.

References

- Chubb, H., Simpson, J.: The use of z-scores in paediatric cardiology. Ann. Pediatric Cardiol. 5, 179–184 (2012). https://doi.org/10.4103/0974-2069.99622
- Daubechies, I.: Orthonormal bases of compactly supported wavelets. Commun. Pure Appl. Math. 41(7), 909–996 (1988). https://doi.org/10.1002/cpa.3160410705. https://onlinelibrary. wiley.com/doi/abs/10.1002/cpa.3160410705
- Gokhroo, R., Anantharaj, A., Bisht, D., Kishor, K., Plakkal, N., Mondal, N.: A pediatric echocardiographic z-score nomogram for a developing country: Indian pediatric echocardiography study—the z-score. Ann. Pediatric Cardiol. 11, 109 (2018). https://doi.org/10.4103/apc. APC_123_17
- Long, M., Lei, W., Mou, L., Zhang, K., Liu, L., Li, Y., Liu, X., YU, W., Gao, G., Chen, X., Shen, W., Shrestha, A.: Z-score transformation of ADC values: a way to universal cut off between malignant and benign lymph nodes. Eur. J. Radiol. **106** (2018). https://doi.org/10.1016/j.ejrad. 2018.07.022
- Melillo, P., Izzo, R., Orrico, A., Scala, P., Attanasio, M., Mirra, M., Luca, N., Pecchia, L.: Automatic prediction of cardiovascular and cerebrovascular events using HRV analysis. PLOS ONE 10, e0118504 (2015). https://doi.org/10.1371/journal.pone.0118504

- Ni, H., Wang, Y., Xu, G., Shao, Z., Zhang, W., Zhou, X.: Multiscale fine-grained heart rate variability analysis for recognizing the severity of hypertension. Comput. Math. Methods Med. 2019, 1–9 (2019). https://doi.org/10.1155/2019/4936179
- Quachtran, B., Hamilton, R., Scalzo, F.: Detection of intracranial hypertension using deep learning. In: Proceedings of the International Conference on Pattern Recognition (IAPR) (2016), pp. 2491–2496. https://doi.org/10.1109/ICPR.2016.7900010
- Rajput, J.S., Sharma, M., Acharya, U.R.: Hypertension diagnosis index for discrimination of high-risk hypertension ECG signals using optimal orthogonal wavelet filter bank. Int. J. Environ. Research and Public Health 16(21) (2019), https://www.mdpi.com/1660-4601/16/ 21/4068
- Rajput, J.S., Sharma, M., Kumbhani, D., Acharya, U.R.: Automated detection of hypertension using wavelet transform and nonlinear techniques with ballistocardiogram signals. Inform. Med. Unlocked 26, 100736 (2021)
- Rajput, J.S., Sharma, M., Tan, R.S., Acharya, U.R.: Automated detection of severity of hypertension ecg signals using an optimal bi-orthogonal wavelet filter bank. Computers Biol. Med. 103924 (2020). https://doi.org/10.1016/j.compbiomed.2020.103924. http://www. sciencedirect.com/science/article/pii/S001048252030264X
- Sakr, S., Elshawi, R., Ahmed, A., Qureshi, W.T., Brawner, C., Keteyian, S., Blaha, M.J., Al-Mallah, M.H.: Using machine learning on cardiorespiratory fitness data for predicting hypertension: the Henry ford exercise testing (fit) project. PLOS ONE 13(4), 1–18 (2018). https://doi.org/10.1371/journal.pone.0195344. https://doi.org/10.1371/journal.pone.0195344
- 12. Sau, A., Bhakta, I.: Screening of anxiety and depression among the seafarers using machine learning technology. Inform. Med. Unlocked (2018)
- Shah, S., Sharma, M., Deb, D., Pachori, R.B.: 2019 International Conference on Machine Intelligence and Signal Analysis Advances in Intelligent Systems and Computing, vol. 748. Springer, Singapore, pp. 473–483 (2019). https://doi.org/10.1007/978-981-13-0923-6_41
- Sharma, M., Acharya, U.R.: Automated detection of schizophrenia using optimal waveletbased l₁ norm features extracted from single-channel EEG. Cogn. Neurodyn. 1–14 (2021). https://doi.org/10.1007/s11571-020-09655-w
- Sharma, M., Bhurane, A., Acharya, U.R.: MMSFL-OWFB: a novel class of orthogonal wavelet filters for epileptic seizure detection. Knowl. Based Syst. (2018). https://doi.org/10.1016/j. knosys.2018.07.019
- Sharma, M., Patel, V., Acharya, U.R.: Automated identification of insomnia using optimal bi-orthogonal wavelet transform technique with single-channel EEG signals. Knowl. Based Syst. 107078 (2021). https://doi.org/10.1016/j.knosys.2021.107078. https://www. sciencedirect.com/science/article/pii/S0950705121003415
- Sharma, M., Pv, A., Pachori, R., Gadre, V.: A parametrization technique to design joint timefrequency optimized discrete-time biorthogonal wavelet bases. Signal Process. 135 (2017). https://doi.org/10.1016/j.sigpro.2016.12.019
- Sharma, M., Rajput, J.S., Tan, R.S., Acharya, U.R.: Automated detection of hypertension using physiological signals: a review. Int. J. Environ. Res. Public Health 18(11) (2021). https://doi. org/10.3390/ijerph18115838. https://www.mdpi.com/1660-4601/18/11/5838
- Sharma, M., Raval, M., Acharya, U.R.: A new approach to identify obstructive sleep apnea using an optimal orthogonal wavelet filter bank with ECG signals (2019). https://doi.org/10. 1016/j.imu.2019.100170
- Sharma, M., Tan, R.S., Acharya, U.R.: A novel automated diagnostic system for classification of myocardial infarction ECG signals using an optimal biorthogonal filter bank. Computers Biol. Med. (2018)
- Sharma, M., Tiwari, J., Acharya, U.R.: Automatic sleep-stage scoring in healthy and sleep disorder patients using optimal wavelet filter bank technique with EEG signals. Int. J. Environ. Res. Public Health 18(6) (2021). https://doi.org/10.3390/ijerph18063087. https://www.mdpi. com/1660-4601/18/6/3087
- Simjanoska, M., Gjoreski, M., Gams, M., Madevska Bogdanova, A.: Non-invasive blood pressure estimation from ECG using machine learning techniques. Sensors 18 (2018). https://doi.org/10.3390/s18041160

- Song, Y., Ni, H., Zhou, X., Zhao, W., Wang, T.: Extracting Features for Cardiovascular Disease Classification Based on Ballistocardiography, pp. 1230–1235 (2015). https://doi.org/10.1109/ UIC-ATC-ScalCom-CBDCom-IoP.2015.223
- Whelton, P., Carey, M.R.: Guideline for the prevention, detection, evaluation, and management of high blood pressure in adults. J. Am. College Cardiol. 71 (2017). https://doi.org/10.1016/j. jacc.2017.11.005
- Xing, W., Bei, Y.: Medical health big data classification based on KNN classification algorithm. IEEE Access 8, 28808–28819 (2020). https://doi.org/10.1109/ACCESS.2019.2955754
- Sharma, M., Darji, J., Thakrar, M., Acharya, U.R.: Automated identification of sleep disorders using wavelet-based features extracted from electrooculogram and electromyogram signals. Computers in Biol. Med. 105224 (2022)
- Sharma, M., Dhiman, H.S., Acharya, U.R.: Automatic identification of insomnia using optimal antisymmetric biorthogonal wavelet filter bank with ECG signals. Computers in Biol. Med. 104246 (2021). https://doi.org/10.1016/j.compbiomed.2021.104246. https://www. sciencedirect.com/science/article/pii/S0010482521000408
- Sharma, M., Kumbhani, D., Yadav, A., Acharya, U.R.: Automated sleep apnea detection using optimal duration-frequency concentrated wavelet-based features of pulse oximetry signals. Appl. Intell. 1–13 (2021)
- 29. Sharma, M., Patel, S., Acharya, U.R.: Automated detection of abnormal EEG signals using localized wavelet filter banks. Pattern Recogn. Lett. (2020)
- Sharma, M., Patel, S., Acharya, U.R.: Expert system for detection of congestive heart failure using optimal wavelet and heart rate variability signals for wireless cloud-based environment. Expert Syst. e12903 (2021)
- Sharma, M., Patel, V., Tiwari, J., Acharya, U.R.: Automated characterization of cyclic alternating pattern using wavelet-based features and ensemble learning techniques with EEG signals. Diagnostics 11(8) (2021). https://doi.org/10.3390/diagnostics11081380. https://www.mdpi. com/2075-4418/11/8/1380
- Sharma, M., Achuth, P., Deb, D., Puthankattil, S.D., Acharya, U.R.: An automated diagnosis of depression using three-channel bandwidth-duration localized wavelet filter bank with EEG signals. Cogn. Syst. Res. 52, 508–520 (2018). http://www.sciencedirect.com/science/article/ pii/S1389041718302298
- Sharma, M., Tan, R.S., Acharya, U.R.: Detection of shockable ventricular arrhythmia using optimal orthogonal wavelet filters. Neural Comput. Appl. (2019). https://doi.org/10.1007/ s00521-019-04061-8
- 34. Sharma, M., Tiwari, J., Patel, V., Acharya, U.R.: Automated identification of sleep disorder types using triplet half-band filter and ensemble machine learning techniques with EEG signals. Electronics 10(13) (2021). https://doi.org/10.3390/electronics10131531