International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Special Issue, 7, March 2019 E-ISSN: 2347-2693

ECG Signal Feature Extraction and Classification: Survey

Bhagyashri Bhirud[1*], Vinod Pachghare[2]

[1,2]Dept. of Computer Engineering, College of Engineering Pune, India

*Corresponding Author: bhirudbp17.is@coep.ac.in, Tel.: +91 9049386143

Available online at: www.ijcseonline.org

***Abstract*—** Nowadays due to busy and hectic lifestyle, many people cannot pay enough attention to their health. Stress, junk food, obesity, smoking and lack of exercise leads to heart diseases. It is one of major cause leading rise to death rate. ECG (Electrocardiogram) is the best and easiest way to record and analyze the electrical and muscular activities of heart. Due to nonlinearity and complexity of ECG signals, it requires significant amount of training to analyze and study the ECG waveform. For preventive measures and predictive analysis, it is necessary to analyze these waveforms in faster, efficient way and in real time too. Number of methods and techniques have been developed in recent time. Different techniques and methodologies are discussed in this literature review.

***Keywords*—** ECG Signal, Arrhythmia, Feature Extraction, Feature Selection

I. Introduction

Due to fast paced life style and lack of nutritious food, many health issues are besieging humans. However, due to these facts, awareness about health, healthy habits and overall healthy lifestyle is increasing among the youngsters. They are cautious about their health. They are opting for different applications or gadgets which help them in monitoring their health. Recently, different electronic devices have been introduced in market that help to continuously monitor the heart rate. To analyze this heart rate data, different techniques and methods have been proposed. There are number of algorithms and number of methods are available for the analysis of ECG signal. Also, different authors refer different features. With all this, the accuracy of the proposed system also plays an important role. Due to this, it is important to review the techniques and methods proposed in recent time, before one can dive into it.

In this paper, we are reviewing the recent work in ECG data analysis. The survey is divided into databases, preprocessing, feature extraction and feature selection, and classification algorithms.

II. Database

Data is the most crucial factor in case of machine learning. It plays an important role in training the model. Different databases are available for every field. For ECG signal data, many datasets are available. It is found that MIT-BIH is the most widely used dataset during ECG signal analysis. Different datasets available are

1. MIT-BIH Dataset
2. AHA Dataset
3. CU Dataset

A. MIT-BIH Dataset

MIT-BIH (Massachusetts Institute of Technology-Beth Israel Hospital). This database contains 48 records approximately 30-min long. Those are obtained from 47 different patients. The records are digitized at 360Hz with 11-bit resolution over 10mV range.

B. AHA Dataset

AHA (American Heart Association) dataset was developed mainly for the evaluation of Ventricular Arrhythmia detectors by American Health Association. In this, 80 two channel records are available sampled at 250Hz with 12-bit resolution over 10mV range. There are 8 classes of 10 recordings under each. These are - no ventricular ectopy, isolated unifocal PVCs, isolated multifocal PVCs, ventricular bi- and trigeminy, R-on-T PVCs, ventricular couplets, ventricular tachycardia and ventricular flutter/fibrillation.

C. CU Dataset

CU (Creighton University) dataset contains 35 recordings 8-min long collected from patients suffering from ventricular tachycardia, ventricular flutter, and ventricular fibrillation. The records are pre-processed using second-order Bessel low-pass filter with cut-off frequency 70Hz. Records are digitized at 250Hz with 12-bit resolution over 10V.

In [1], they have used MIT-BIH Arrhythmia database. Along with this, they have utilized real time data captured by NI

(National Instruments) DAQ (Data Acquisition) [6]. While in [2], they have used three databases. CU Ventricular Tachyarrhythmia database, MIT-BIH Malignant Ventricular Arrhythmia database and AHA database. They have considered all the following signals as non-VF (Ventricular Fibrillation) – Normal sinus, atrial fibrillation, ventricular bigeminy, first-degree heart block, nodal rhythm, sinus bradycardia, paced rhythm, supraventricular tachyarrhythmia, high-grade ventricular ectopic activity and ventricular escape. They excluded signals labelled as noise. Authors of [3] have used MIT-BIH Arrhythmia dataset. Also, [4] have used three databases. Medical imaging Technology database (MITDB), Creighton University ventricular tachycardia (CUDB), MIT-BIH malignant ventricular arrhythmia database (VFDB). Two databases are used in [5]. First is MIT-BIH arrhythmia database and second is arrhythmia data obtained by wearable device, as they have proposed method for automatic wearable electrocardiogram classification.

III. PREPROCESSING

The databases available has baseline wander, motion artifacts, power interference. Different authors used different methods for denoising of ECG signal.

According to [1], datasets has baseline wander, motion artifacts and white Gaussian noise. For removal of white Gaussian noise, they have used Wiener filtering. Wiener filter is not adaptive filter as it considers input to be stationery. It removes the noise by comparing the input signal with the desired noise free signal. There are several options of adaptive filtering based on wavelet like EMD (Empirical Mode Decomposition), RLS (Recursive Least Squares) and LMS (Least Mean Square). As EMD and RLS require more computational power, they have used LMS adaptive filtering for noise removal. However, LMS filter comes with fixed step size. So, they suggest normalized step size. They are using a delayed error normalized LMS.

In [2], they have re-sampled the data at 250Hz sampling frequency. High-pass filter is used with 1 Hz cut off frequency for removal of baseline wander. They have used second-order 30 Hz low-pass filter for high-frequency noise. Also, a notch filter is used to eliminate power interference.

In [3], they used a bandpass (Butterworth) filter with 0.5Hz and 40Hz cut off frequencies for noise removal. To reduce baseline drift to zero, they referred method mentioned in "Total removal of baseline drift from ECG signal" [7].

According to [4], four steps are required to remove noise from ECG signals.

1. Mean substitution
2. Five order moving average filtering
3. High pass filtering with $f_c = 1\,Hz$ (for drift suppression)
4. Low pass Butterworth filtering with $f_c = 30\,Hz$

In [5], initially P wave and QRS complex are removed using a 200 ms width median filter and T wave is removed by 600ms width median filter. Then power line and high frequency noise is removed by a 12-order low-pass filter with 35Hz cut-off frequency.

IV. FEATURE EXTRACTION

Feature extraction plays a very important role in the classification. The performance of the classification algorithm highly depends upon the features extracted from the data. This section covers the popular methods of feature extraction and selection. Sometimes, feature extraction and selection terms are used simultaneously. However, feature extraction is extracting the information from the data available. Feature selection is process of selecting the most useful and relevant features out of all in order to efficiently classify the data.

A. Features

Different authors considered different features, for which different methods are used for extraction.

In [1], total 14 features are considered, 9 time-domain features and 6 frequency domain features.

- Time domain features - RR mean(ms), RR STD (ms), HR mean(bpm), HR STD (bpm), RMSSD (ms), NN50, pNN50, RR triangular index, TINN (ms)
- Frequency domain features - VLF power (ms2), LF power (ms2), HF power (ms2), LF norm, HF norm, LF/ HF ratio

In [2], total 14 metrics are extracted under different categories.

1. Complexity measure ($Complexity$) [9] - ECG signal is converted into a binary signal. A 0/1 binary string is generated by comparing ECG data to a threshold. Complexity measure is computed using the following
   a. At beginning, c(n) = 1, $S = s_1$, Q = s_2, $SQ = s_1$
   b. After number of iterations, $S = s_1, s_2, s_3 \dots . s_r$, $Q = s_{r+1}$
   c. If Q is a substring of SQ_p, S is unchanged and Q becomes $Q = s_{r+1}, s_{r+2}$ until Q is not a substring of SQ_p
   d. Then S is renewed by combining it with Q. $S = s_1, s_2, s_3 \dots . s_r, s_{r+i}$, $Q = s_{r+i+1}$ and $c(n) = c(n) + 1$
2. VF-filter Leakage measure ($Leakage$) [10] - A narrow bandstop filter with central frequency

equivalent to mean signal frequency. The output is VF-filter leakage.

3. Spectral analysis [11] - Data segment multiplied with Hamming window and transform to a frequency domain by Fast Fourier Transform (FFT). Four spectral parameters are used -$FSMN$, $A1$, $A2$, $A3$.
4. Time delay algorithm ($Timedelay$) [12] - Signal $s(t)$ is plotted by $x(t)$ on x-axis, $x(t+T)$ on y-axis, T – time constant. In this paper, $T = 0.5s$ is used for VF detection. ECG signal is down-sampled to 50Hz frequency and phase space plot is plotted on 40x40 grid. Area of the plot covered by the curve is counted.
$$Timedelay = \frac{numberOfVisitedBoxes}{numberOfAllBoxes}$$
5. Bandpass filter and auxiliary counts [13] - ECG data is filtered with an integer coefficient recursive digital filter with central frequency at 14.6Hz and bandwidth from 13-16.5Hz. The sampling frequency is 250 Hz. Three auxiliary parameters are calculated from filter signal - $count1$, $count2$ and $count3$. These represent the number of signal samples with amplitude values within a certain range.
6. Covariance calculation ($CovarBin$) [14] - measures the variance of the corresponding binary signal of ECG.
7. Frequency calculation ($FreqBin$) [14] - calculated by counting the number of binary signal transitions and dividing by window length.
8. Area calculation ($AreaBin$) [14] - realized by summing the values of the binary signal samples.
9. Kurtosis ($kurtosis$) [15] - It is the fourth standardized moment of ECG.

In [3], 13 features are extracted. These are - P-peak, Q-peak, R-peak, S-peak, T-peak, the time duration of P wave, the time duration of Q wave, the time duration of R wave, the time duration of s wave, time duration of T wave, P-R interval, R-R interval, S-T interval.

In [4], ECG signal is divided into 8s segments. For these segments, 13 parameters were identified falling in three major categories

1. Temporal/Morphological parameters - these parameters are defined in the time domain
   a. Threshold crossing Interval (TCI) [16]
   b. Threshold crossing sample count ($TCSC$) [17]
   c. Standard exponential (STE) [18]
   d. Modified exponential (MEA) [18]
   e. Mean absolute value (MAV) [19]
2. Spectral parameters – these parameters are defined in the frequency domain
   a. VF filter ($VFleak$) [10]
   b. Spectral algorithm (M) [11]
   c. Spectral algorithm ($A2$) [11]
   d. Medium frequency (FM) [20]
3. Complexity parameters – these parameters measure the complexity of the ECG signal
   a. Complexity measurement (CM) [21]
   b. Phase space reconstruction (PSR) [12]
   c. Hilbert transform ($HILB$) [22]
   d. Sample entropy ($SpEn$) [23]

In [5], for feature representation of ECG data, ECG morphology features and two temporal features are used. For extracting these features, QRS is detected using *ecgpuwave* software available on physionet.org website [8].

1. Morphological features are selected at 100ms before R-peak and 450ms after R-peak.
2. Temporal features are
   a. Pre-RR interval - distance between current R-peak and its previous R-peak
   b. Post-RR interval - distance between current R-peak and next R-peak

B. Feature Extraction Methods

In [1], Discrete Wavelet Transform method is used on denoised ECG to extract R-peaks for HRV (High Rate Variability) features. They have used Coiflet wavelet for HRV feature extraction over Daubechies and Spline wavelets for R-peak extraction. Frequency domain analysis of HRV signal helps extract more features such as VLF, LF, HF.

In [2], their algorithm works in two phases: Development phase and validation phase. In development phase, the VF metrics are extracted from specific length ECG record based on recent studies.

According to [3], their proposed classifier totally depends upon the feature extracted in extraction phase. Modified Pan Tompkins algorithm is applied to filtered ECG signal, to determine R-peaks. First high slope is obtained using differentiation equation, then the signal is squared to detect R-peaks and finally, integration sum is performed to extract the slope of R wave. Once R-peaks are extracted, these are used as a reference to extract other peaks (P, Q, S and T).

In [4], for 8-s segment of ECG signal, they have computed 13 previously defined parameters.

In [5], Stacked Denoising Autoencoder (SDAE) is used to learn the feature representation. Being an unsupervised algorithm and a neural network, autoencoder consists of three layers – visible layer, hidden layer and reconstruction layer. The encoding part of autoencoder maps input to hidden representation using activation function. Here, they have used sigmoid as activation function. Decoding part again reconstructs the input vector from hidden representation. Hidden nodes are imposed with sparsity constraint that helps in identifying interesting structure in data. Sparsity constraint r is a small positive value close to zero.

C. Feature Selection Methods

Feature selection is the process to be performed after feature extraction. It helps to select the most relevant features to efficiently classify the ECG signal.

According to [1], all 14 features extracted are fed to the classifier. They haven't used any feature selection mechanism.

In [2], they have used Genetic Algorithm (GA) as a feature selection technique. Every individual VF metric was fed to SVM classifier. According to performance of each, GA was used to select optimal subset of VF metrics. A 14-element chromosome is defined (as VF metrics are 14). The details of GA – 50 chromosome population, 5% mutation rate, 5% cloning rate and 95% cull rate for crossover and 50 generation limits. After repeating this for 150 times, the variables were sorted and selected according to the frequency of selection. At the end, 9 metrics were selected. [$Leakage$, $Count2$, $CovarBin$, $FreqBin$, $AreaBin$, $Kurtiosis$, $Complexity$, $A3$ and $Count$].

In [3], they have extracted 13 features and all of them are used for classification.

In [4], the filter-type of feature selection method is utilized. In filter-based methods, features are ranked according to some criteria and the ones with the lowest score are omitted. They have used a combination of Correlation criteria, maximum separability Fisher criteria and mutual information method [mRMR minimal redundancy maximal relevance]. Combining the scores of these three, features are ranked.

V. CLASSIFICATION TECHNIQUES

Once the features are extracted from the data and optimal subset is selected, then those are fed as input for the classification algorithm. Based on these features, classification of the ECG signal can be performed. Number of algorithms are available for the training of the model such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), Deep Neural Networks (DNN), Random Forests, K-Nearest Neighbor (KNN) and Bayesian networks. However, it is found that SVM is the most widely used classifier.

In [1], SVM classifier is used to classify the ECG signals into normal and abnormal classes. For better performance, cost function is used with different weights. 90% of data, signals of 60s length, are used for training of SVM classifier and 10% is used to validate the accuracy of classifier. Maximum classification accuracy of 96% is achieved with the work in this paper.

In [2], they have also used SVM as a classifier. However, they have used it with a Gaussian radial basis function kernel. In development phase, SVM model was trained with a variable number of ECG parameters as input with every possible combination. Initially, individual parameters, then in pairs, then in triplets and so on. The ones giving best performance were selected for each number of features. In validation phase, each selected feature combination was validated using five-fold cross-validation. The five-fold procedure was repeated for 50 times and the average was used for performance. This whole process was repeated on different window lengths.

The performance measures used in [2] are Sensitivity (Se), Specificity (Sp), Accuracy (Ac) and Area under the curve (AUC). The best performance was observed when pair of parameters (Count2 and Leakage) were used with accuracy of 96.4%. It was observed that accuracy rate is lower when window length is shorter and balancing the data does not add any value in increasing the performance. The Se for VF detection increases with number of features while SP decreases. As compared to 10s window, use of 5s window does not substantially modifies the result. Their method reports both in-sample training data results and out-of-sample validation data results. As the standard databases are used, it is suggested that data with high noise can be examined. Also, exhaustive set of features can be studied.

In [3], they have proposed the classifier based on the diagnosis of cardiologist. They have considered the previous normal values of different features of ECG and based on that the classifier is proposed. Table 1 shows the normal values for different features.

Table 1. Characteristics of Normal Waves and Intervals

Waves and intervals	**Duration**	**Amplitude**	**Comments**
P-wave	< 0.11s	< 2.5mm	Must be upright and followed by a QRS complex
QRS-complex	< 0.12s	> 0.5mV	Upright R and R-wave amplitude < 20mm
T-wave	N/A	N/A	Always upright
P-R interval	0.12 – 0.22s	N/A	-
S-T interval	N/A	> 0.5mm	Isoelectric, slanting upwards to the T-wave
Q-T interval	< 0.45s	N/A	-

The performance metrics used in [3] are Sensitivity (Se), Positive predictivity ($+P$), Detection error rate (DER) and Accuracy (Acc). The author of this paper compared the result of proposed classifier with existing classifiers. Two neural network classifiers [Feed Forward Neural Network (FFNN) and Multilayered Perceptron (MLP)], four SVM classifiers [Linear-SVM, Radial Basis Function (RBF), Polynomial-SVM and Quadratic-SVM] and K Nearest Neighbor (KNN) are used for comparison. Different proportions of training and test data were used – 95% 5%, 85% 15% and 75% 25%. The overall result shows Se – 99.98% and $+P$ of 99.9% with DER of 0.018%. Overall accuracy is 99% with average computational time 0.006203s. In some cases, MLP performs better than proposed classifier. However, its computation time is more. As the proposed classifier is dependent on the features extracted, it is sensitive to quality of ECG signal.

In [4], SVM classifier with Gaussian kernel is used. Bootstrap re-sampling is used for the estimation of the performance of SVM classifier with $R = 500$. First individual parameter was assessed to correctly classify VF vs non-VF and shockable versus non-shockable signals. Receiver operating characteristic curve (ROC) is analyzed for each parameter. Area under curve (AUC), Se and Sp are used as performance metrics. $TCSC$, $SpEn$ and $VFleak$ showed high values of Se and Sp. For performance of SVM with all parameters is evaluated with 70% training and 30% testing data. Balanced error rate (BER) was used as metric to set parameters of SVM [C, g]. Metrics used for SVM performance are –Se, Sp, AUC, positive predictivity ($+P$) and accuracy (Acc). Overall performance – Se – 93.6%, Sp – 98.8%, $+P$ – 88.3% Acc – 98.4%, BER – 3.7, AUC – 99.4%.

In [5], SoftMax regression is used for classification. It is supervised learning algorithm which is added on resulting hidden representation layer. Active learning is used to finetune DNN and to reduce need for large amounts of labeled samples. Active learning corresponds to labeling new unlabeled samples and adding most relevant to training set. For this, DNN posterior probabilities are used to associate confidence measure. Two methods are used based on posterior probabilities – Breaking ties (BT) algorithm and Modified breaking ties (MBT) algorithm. To get the posterior probability, SVM is trained according to Platt's method.

Architecture of DNN used in [5] - Architecture of encoding part - 52 nodes in visible layer and 200 nodes in hidden layer. Learning rate is 1, epoch is 400, batch size is 10, denoising parameter is 0.5. Classification is performed according to AAMI standard. Classification is evaluated using Accuracy (Acc), Sensitivity (Se), Specificity (Sp) and Positive predictivity (Ppr).

VI. CONCLUSIONS

In the literature, different algorithms and classifiers used for ECG analysis are discussed. Few researchers had focused on the features and the methods to extract those. On the other hand, some researchers explored and developed classifiers with minimal work on features. Most of the researchers discussed in literature have used MIT-BIH dataset, which is unbalanced. This factor is not considered by many researchers. Several methods only provide binary classification – normal and abnormal. The methods, providing classification with a greater number of classes, tend to increase the complexity. Also, machine learning approaches depends most upon the data that is used to train the model, rather that the choice of algorithm. This leads to focus on choosing or making available ones the quality dataset in future works.

Now new era is emerging with remote healthcare and mobile health technology. So, there is a need for development of fast and accurate algorithms with less complexity to be run in lightweight end devices.

REFERENCES

[1] Venkatesan, C., P. Karthigaikumar, Anand Paul, S. Satheeskumaran, and R. Kumar. "ECG Signal Preprocessing and SVM Classifier-Based Abnormality Detection in Remote Healthcare Applications." IEEE Access 6 (2018): 9767-9773

[2] Li, Qiao, Cadathur Rajagopalan, and Gari D. Clifford. "Ventricular fibrillation and tachycardia classification using a machine learning approach." IEEE Transactions on Biomedical Engineering 61, no. 6 (2014): 1607-1613

[3] Hammad, Mohamed, Asmaa Maher, Kuanquan Wang, Feng Jiang, and Moussa Amrani. "Detection of abnormal heart conditions based on characteristics of ECG signals." Measurement 125 (2018): 634-644

[4] Alonso-Atienza, Felipe, Eduardo Morgado, Lorena Fernandez-Martinez, Arcadi Garc´ıa-Alberola, and Jos´e Luis Rojo-Alvarez. "Detection of life-threatening arrhythmias using feature selection and support vector machines." IEEE Trans. Biomed. Eng 61, no. 3 (2014): 832-840

[5] Xia, Yufa, Huailing Zhang, Lin Xu, Zhifan Gao, Heye Zhang, Huafeng Liu, and Shuo Li. "An automatic cardiac arrhythmia classification system with wearable electrocardiogram." IEEE Access 6 (2018): 16529-16538

[6] http://www.ni.com/data-acquisition/what-is/

[7] Chouhan, V. S., and Sarabjeet Singh Mehta. "Total removal of baseline drift from ECG signal." In Computing: Theory and Applications, 2007. ICCTA'07. International Conference on, pp. 512-515. IEEE, 2007.

[8] https://www.physionet.org/physiotools/wag/ecgpuw-1.htm

[9] Zhang, Xu-Sheng, Yi-Sheng Zhu, Nitish V. Thakor, and Zhi-Zhong Wang. "Detecting ventricular tachycardia and fibrillation by complexity measure." IEEE Transactions on biomedical engineering 46, no. 5 (1999): 548-555

[10] Kuo, S. "Computer detection of ventricular fibrillation." Proc. of Computers in Cardiology, IEEE Comupter Society (1978): 347-349

[11] Barro, S., R. Ruiz, D. Cabello, and J. Mira. ”Algorithmic sequential decision-making in the frequency domain for life threatening ventricular arrhythmias and imitative artefacts: a diagnostic system.” Journal of biomedical engineering 11, no. 4 (1989): 320-328

[12] Amann, Anton, Robert Tratnig, and Karl Unterkofler. ”Detecting ventricular fibrillation by time-delay methods.” IEEE Transactions on Biomedical Engineering 54, no. 1 (2007): 174-177

[13] Jekova, Irena, and Vessela Krasteva. ”Real time detection of ventricular fibrillation and tachycardia.” Physiological measurement 25, no. 5 (2004): 1167

[14] Jekova, Irena. ”Shock advisory tool: Detection of life-threatening cardiac arrhythmias and shock success prediction by means of a common parameter set.” Biomedical Signal Processing and Control 2, no. 1 (2007): 25-33

[15] Li, Qiao, Roger G. Mark, and Gari D. Clifford. ”Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter.” Physiological measurement 29, no. 1 (2007): 15

[16] Thakor, Nitish V., Y-S. Zhu, and K-Y. Pan. ”Ventricular tachycardia and fibrillation detection by a sequential hypothesis testing algorithm.” IEEE Transactions on Biomedical Engineering 37, no. 9 (1990): 837-843

[17] Arafat, Muhammad Abdullah, Abdul Wadud Chowdhury, and Md Kamrul Hasan. ”A simple time domain algorithm for the detection of ventricular fibrillation in electrocardiogram.” Signal, Image and Video Processing 5, no. 1 (2011): 1-10

[18] Amann, Anton, Robert Tratnig, and Karl Unterkofler. ”Reliability of old and new ventricular fibrillation detection algorithms for automated external defibrillators.” Biomedical engineering online 4, no. 1 (2005): 60

[19] Anas, Emran M. Abu, Soo Y. Lee, and Md K. Hasan. ”Sequential algorithm for life threatening cardiac pathologies detection based on mean signal strength and EMD functions.” Biomedical engineering online 9, no. 1 (2010): 43

[20] Dzwonczyk, Roger, Charles G. Brown, and Howard A. Werman. ”The median frequency of the ECG during ventricular fibrillation: its use in an algorithm for estimating the duration of cardiac arrest.” IEEE Transactions on Biomedical Engineering 37, no. 6 (1990): 640-646

[21] Zhang, Xu-Sheng, Yi-Sheng Zhu, Nitish V. Thakor, and Zhi-Zhong Wang. ”Detecting ventricular tachycardia and fibrillation by complexity measure.” IEEE Transactions on biomedical engineering 46, no. 5 (1999): 548-555

[22] Amann, Anton, Robert Tratnig, and Karl Unterkofler. ”A new ventricular fibrillation detection algorithm for automated external defibrillators.” database 1, no. 2 (2005): 3

[23] Li, Haiyan, Wenguang Han, Chao Hu, and Max Q-H. Meng. ”Detecting ventricular fibrillation by fast algorithm of dynamic sample entropy.” In Robotics and Biomimetics (ROBIO), 2009 IEEE International Conference on, pp. 1105-1110. IEEE, 2009

Authors Profile

Ms. B P Bhirud pursed Bachelor of Technology in Information Technology specialization from Walchand College of Engineering, Sangli in 2015. She is currently pursuing Master of Technology in Information Security from College of Engineering, Pune. She is working in the domain of Data analytics and machine learning for remote health care systems.

Dr V K Pachghare Ph.D., is Associate Professor in the Department of Computer Engineering and Information Technology, College of Engineering, Pune (An autonomous institute of Government of Maharashtra, India). He has 25 years of teaching experience and has published books on Cryptography and Information Security (2nd edition, 2015) and Computer Graphics. He has over 30 research publications in various international journals and conferences. He has been recently honored with Best Faculty of the year 2018 Award by CSI Mumbai chapter in TextNext 2018 at IIT Bombay.

