International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-10 Issue-2, December 2020

Ecg Heartbeat Classification: Conceptual Understanding through Cnn & Rnn – A Machine Learning Approach

P. Rama Santosh Naidu, G. Lavanya Devi, Kondapalli Venkata Ramana

***Abstract:** In recent days Machine Learning has become major study aspect in various applications that includes medical care where convenient discovery of anomalies in ECG signals plays an important role in monitoring patient's condition regularly. This study concentrates on various MachineLearning techniques applied for classification of ECG signals which include CNN and RNN. In the past few years, it is being observed that CNN is playing a dominant role in feature extraction from which we can infer that machine learning techniques have been showing accuracy and progress in classification of ECG signals. Therefore, this paper includes Convolutional Neural Network and Recurrent Neural Network which is being classified into two types for better results from considerably increased depth.*

***Key words:** Basic CNN, DeepResidual CNN, Convolution layer, Max pool block*

I. INTRODUCTION

We visualize 5000 sampled data points using the embeddings achieved by our baseline CNNs.

Here we show the visualization obtained from applying PCA to extract the 2 most informative dimensions from these embeddings.

We also use K-means to preform clustering, showing the results in the same representation.

There is a nice separation of classes in the MITBIH dataset, with some overlap between "Normal" and "S" with the clustering being able to clearly identify the "Q" and "V" classes.

There is some confusion in the "Normal" class, perhaps due to the overwhelming majority of the dataset that the class represents.

For the PTBDB dataset, the separation is not as convincing, with a large overlap in the visualization and with a complete failure in the clustering.

This indicates that the PTBDB dataset may be more challenging to classify.

Further methods, including T-sne, and work on the raw data.

For MIT-BIH dataset:

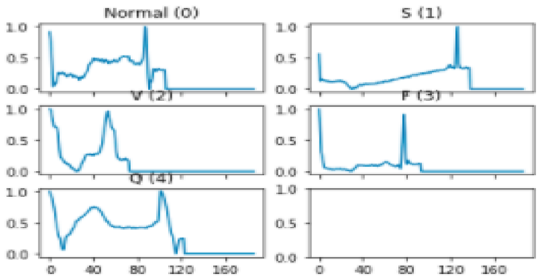

Figure 1: Plot One of Each Label - MITBIH Dataset 'N': NORMAL, 'S': supraventricular ectopic beat, 'V': ventricular ectopic beat, 'F': fusion beat, 'Q': unknown beat

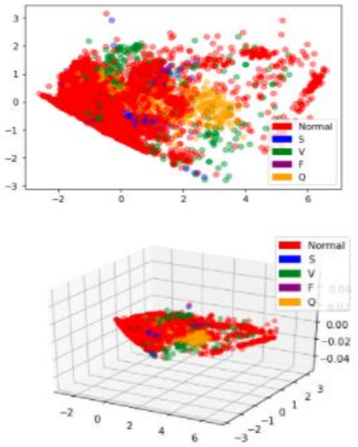

Figure 2: PCA on the Raw Data - MITBIH Dataset PCA on the raw data does not reveal a very separable structure. This suggests simplet linear models would most likely not perform well on the raw data

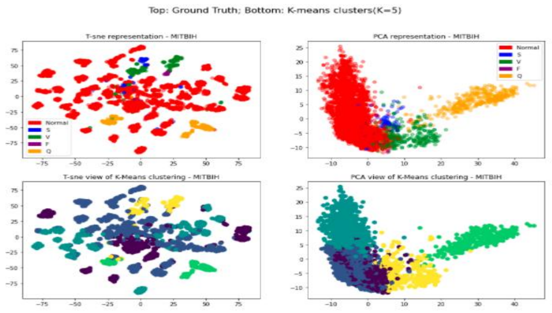

Figure 3: Attempt to Cluster with K-means on the Embeddings - MITBIH Dataset

Revised Manuscript Received on December 30, 2020.
***** Correspondence Author

P. Rama Santosh Naidu*, Assistant Professor in Andhra University College of Engineering (A), Andhra Pradesh, India.

Dr. G. Lavanya Devi, Assistant Professor Andhra University College of Engineering (A) Department of CS & SE, Andhra Pradesh, India.

Dr. K.Venkata Ramana, Assistant Professor in Andhra University College of Engineering (A), Andhra Pradesh, India.

For PTB dataset:

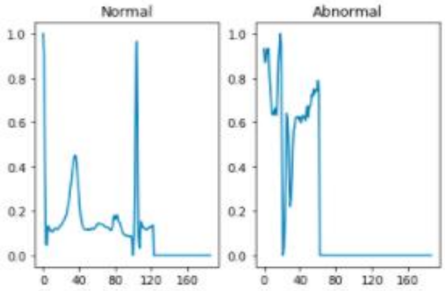

Figure 4: Normal vs Abnormal Count
Once more, quite unbalanced dataset, with abnormal having more than double the datapoints of normal

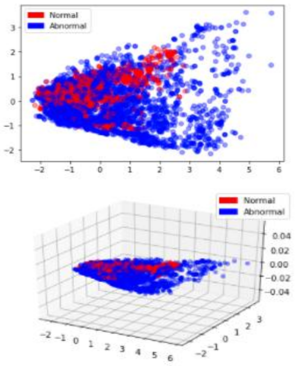

Figure 5: PCA on the Raw Data - PTBDB Dataset
PCA does not show such a clear separation

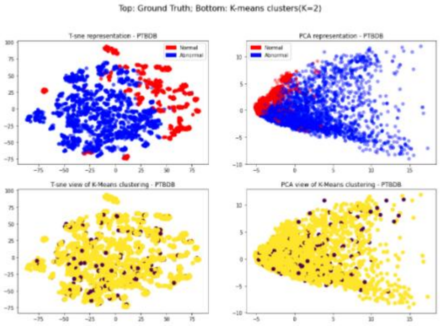

Figure 6: Attempt to Cluster with K-means on the Embeddings - PTBDB Dataset

II. METHODOLOGY

1.1 SAMPLES VISUALIZATION AND CLUSTERING

We used the TSNE algorithm to reduce the number of dimensions from 187 to 2 in both data sets.

Due to the significant difference in class sizes of the MIT dataset we sampled 600 signals per class at random which is roughly the size of the smallest class.

The 2D plots of the MIT and PTB datasets are shown in Figure 1a and 1b accordingly.

For the PTB data set the separation between the normal and abnormal cases is really noticeable.

For the MIT data set, some classes are clearly separated from each other but there are some classes with conflicting points.

This is indicative of the results.

When comparing F1 scores achieved for each dataset it is seen that it is generally 7-8% higher.

For MIT-BIH dataset:

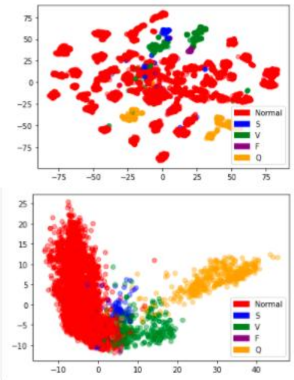

Figure 7: t-SNE on the Embeddings - MITBIH Dataset t-SNE on the embeddings shows a clear separation between classes, even for the rare F class the representation obtained from the baseline CNN reveals a much more separable structure.

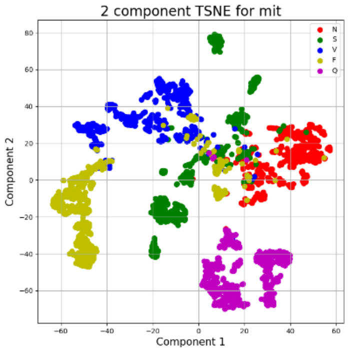

Figure 8: t-SNE applied on MIT data set

For PTB dataset:

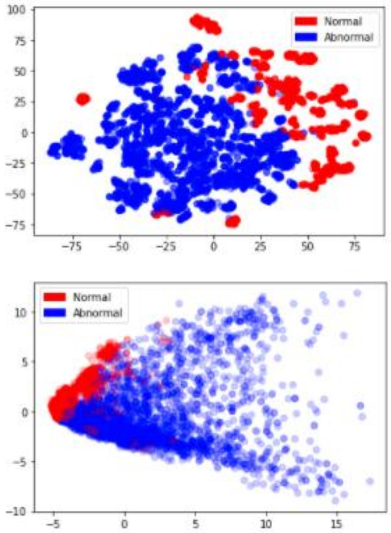

Figure 9: t-SNE on the Embeddings - PTBDB Dataset
But t-sne shows a much clearer separation between both classes. The PCA plot is not as clearly separable, with considerable overlap of both classes

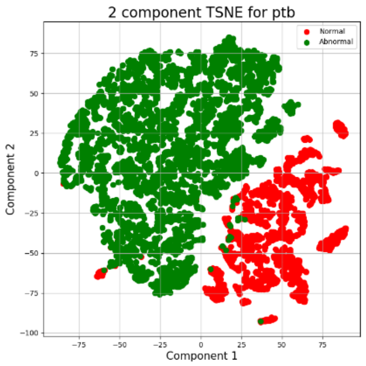

Figure 10: t-SNE applied on PTB data set

2.2 SOLVE BOTH DATASETS WITH RNNS

To solve both datasets using RNNs, we tried different recurrent gate types, and we found that a SimpleRNN works well for both datasets. Both models have two RNN layers and one last Dense layer. We used more hidden units for MIT than for PTB (a good engineering rule seemed to be “total number of parameters = numbers of datapoints”).

For MIT, we got an accuracy of 95.77% and for PTB an accuracy of 87.36%, an AUROC of 83.49% and an AUPRC of 89.05%. These accuracies are lower than the baselines. This is probably because a Conv based net- work is able to use information from both past and future timesteps, whereas an RNN based network only uses past timesteps. This intuition is also confirmed by the higher score we got with Bidirectional RNNs.

2.3 COMPARISON WITH OTHER MODELS

2.3.1 RNN:

For the main task of the project, we developed RNN models for both datasets. The main results from our experiments are that CNNs are a better solution to this task: they are both easier to train and achieve better performance (see comparison table in next section). We found that GRU cells outperform LSTM cells in the recurrent layers, as all LSTM architectures we tested ended up suffering from vanishing gradients and not learning anything. It’s worth noting that the PTBDB task is particularly hard for an RNN because of the small number of samples. In fact, all RNN architectures we tested got stuck on predicting only the dominant class, and we only achieved decent results when we made the network bidirectional.The best hyperparameter values and architectures were obtained through hyperparameter tuning.

2.3.2 BIDIRECTIONAL RNN:

We implemented bidirectional RNNs to use the entire signal (past and future data) to make predictions. We got higher scores than with RNN which strengthens our intuition that using information about the whole signal generally improves prediction. Even with a simple model of two bidirectional layers followed by a dense layer gives an accuracy of 98.51% and 90.18% for the MIT and PTB data sets respectively. We also get an AUROC of 89.85% and AUPRC of 93.38% for PTB. The results are not far from the baselines.

2.3.3 CNN WITH RESIDUAL BLOCKS:

We replicated the algorithm implemented by the authors of [1]. In the article residual blocks were used in a five-layer deep CNN. The code from Kaggle as produced by N. Mine [2] was used as a starting point to replicate the articles algorithm and results. The input and output layers must have the same dimensions. We set the padding to “Causal” and “Same” for the aforementioned reason. Causal padding implies that the output at “t” does depends on the input at “t+1”. The model is similar in most ways to the CNN baseline model with the exception of the residual blocks. Indeed, this method seems to give the best results. For the MIT dataset - “causal” padding: F1 of 93.06%, Accuracy of 98.81%, “same” padding: F1 of 92.85%, Accuracy of 98.78%. For the PTB dataset - “causal” padding: F1 of 99.01%, Accuracy of 99.21%, AUROC of 98.94%, AUPRC of 99.15%, “same” padding: F1 of 99.4%, Accuracy of 99.52%, AUROC of 99.36% and AUPRC of 99.54%.

2.3.4 SUPPORT VECTOR MACHINE (SVM):

We tried two different approaches to the application of SVMs to the dataset. We first applied an SVM on the whole dataset directly which is equivalent to applying an SVM to a 187-dimension feature matrix. Each signal is interpreted by the SVM as a point in 187-dimensional space. Indeed, this approach seems to outperform the handpicked features described below, but took eight times longer. The results for the MIT dataset are - F1 of 95.17% and Accuracy of 94.4%.The results for the PTB dataset are - F1 of 95.77%, Accuracy of 94.02%, AUROC of 94.26% and AUPRC of 96.3%. The second method, was to perform a feature extraction on the given dataset. We identified features that were relevant to the classes. Features that are of medical interest and features that were found to differentiate well the classes were extracted. Usually when medical professionals are called to assess an ECG, they have multiple consecutive heartbeats such that they can calculate features such as the variance in heartbeatrate (RR-peak interval) as an indication of arrhythmia. Our dataset was formed of single heartbeats and hence limited our choice of features. The features included the positions and height of the peaks, the total length of the signal, the gradient at various relevant locations as well as the proportion of the signal above and below a given threshold. These were tuned to better differentiate between the signals. The results for the MIT dataset are - F1 of 83.03% and Accuracy of 77.9%.The results for the PTB dataset are - F1 of 86.15%, Accuracy of 81.76%, AUROC of 84.32% and AUPRC of 90.41%. The soft-margin SVM selected implements an RBF kernel whose hyperparameters were calibrated using a grid-search. To account for the imbalance in the classes we used the in-built Sklearn option for balanced weighing of the classes. We furthermore fitted an SVM as the classifier of the last convolutional layer of the baseline code. This proved difficult to tune for both good precision and recall in the given time-frame.

2.4 ENSEMBLE OF MODELS

We used ensemble methods to increase the performance of our predictions. Two methods were implemented to combine model predictions:

1. the mode function to take the most commonly predicted class over the combined models
2. the SoftMax function to take the class with highest overall probability over the combined models

2.5 TRANSFER LEARNING:

The motivation for transfer learning in this setting is the lack of data: as we discussed previously, a regular RNN network with LSTM or

GRU cells was not able to learn anything meaningful on the small PTBDB dataset. As both datasets consist of ECG data, it could be useful to transfer the "knowledge" from the larger dataset to the prediction task on the smaller dataset. In Figure 3, we show that even though they were trained on another prediction task, the embeddings created by a network on the MITBIH dataset are still useful for discriminating the classes in the PTBDB dataset. This is even more evident when compared to the plain RNN embeddings (Figure 11, bottom right corner), which show no separation power at all. The evaluation results are in agreement with these visualizations: a simple RNN trained on top of MITBIH embeddings is significantly better than one trained from scratch.

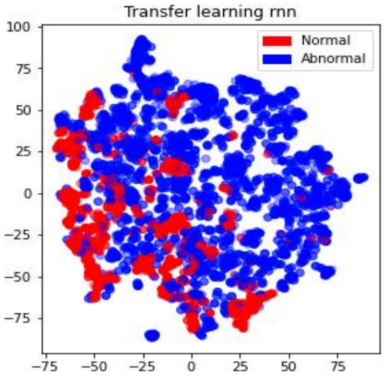

Figure 11: Transfer learning RNN

As suggested, there are different types of transfer learning.

1. **Transfer 1: Transfer learning with RNNs, frozen base model.**
2. **Transfer 2: Transfer learning with RNNs, retraining whole model.**
3. **Transfer 3: Transfer learning with RNNs, first frozen base model, then re-training whole model.**

Results of Transfer1, Transfer2 and Transfer3 are given in below ANALYSIS AND RESULTS section.

III. ANALYSIS AND RESULTS

METHOD	MODE	SOFTMAX
F1 SCORE	98.61%	98.64%
ACCURACY	98.67%	98.69%

Table 1: Results for the ensemble model {RNN + RNN bidirectional + CNN with residual blocks} for the MIT data set

METHOD	MODE	SOFTMAX
F1 SCORE	96.57%	98.06%
ACCURACY	95.07%	97.20%
AUROC	94.32%	96.51%
AUPRC	96.14%	97.56%

Table 2: Results for the ensemble model {RNN + RNN bidirectional + CNN with residual blocks} for the PTB data set

Table 1 and 2 shows that the results are in any cases better than the one obtained using a single model. We get approximately the same scores as for the baseline.

SCORE	TRANSFER 1	TRANSFER 2	TRANSFER 3
F1 SCORE	69.21%	87.42%	89.1%
ACCURACY	78.36%	89.9%	91.24%

Table 3: Comparison of scores for different types of transfer learning

Summary of Results

METHOD	ACCURACY	F1 SCORE
BASELINE	98.31%	91.07%
RESIDUAL BLOCKS DEEP RESIDUAL BLOCKS	98.78% 98.81%	92.85% 93.06%
RNN RNN BIDIRECTIONAL	95.77% 98.51%	77.53% 91.54%
SVM (feat extract) SVM (original data)	77.90% 94.40%	83.03% 95.17%

Table 4: Comparison of all methods for MIT dataset

METHOD	ACCURACY	F1 SCORE	AUROC	AUPRC
BASELINE	99.35%	99.55%	99.09%	99.34%
RESIDUAL BLOCKS DEEP RESIDUAL BLOCKS	99.51% 99.20%	99.40% 99.01%	99.36% 98.84%	99.54% 99.15%
RNN RNN BIDIRECTIONAL	87.36% 90.18%	91.33% 93.01%	83.49% 89.85%	89.05% 83.38%
SVM (feat extract) SVM (original data)	81.76% 94.02%	86.15% 95.77%	84.33% 94.26%	90.41% 96.30%

Table 5: Comparison of all methods for PTB dataset

SCORE	Transfer 1: Transfer learning with RNNs, frozen base model.	Transfer 2: Transfer learning with RNNs, retraining whole model.	Transfer 3: Transfer learning with RNNs, first frozen base model, then re-training whole model.
F1 SCORE	69.21%	87.42%	89.1%
ACCURACY	78.36%	89.9%	91.24%

Table 6: Comparison of scores for different types of transfer learning

IV. CONCLUSION

This paper is useful as the developed CNN model is trained for the selected for MIT-BIH arrhythmia and PTBDB datasets on which the classification is performed. Hence the complexity of the proposed model is less compared to existing models.The viability, accuracy and capacities of our method in classification of ECG signals is exhibited and quantitative examinations with various RNN models have additionally been carried.

The proposed framework can likewise be upgraded to be digital cardiologist for the easy identification of heart conditions.

REFERENCES

1. A Generic Approach of ECG Signals Classification Using CNN by P.Rama Santosh Naidu &Dr.G.Lavanyadevi, Helix (2020) 10 (5): 75-80
2. An Optimized Solution: Deep Residual Learning for ECG Signal Classification by P.Rama Santosh Naidu &Dr.G.Lavanyadevi, Helix (2020) 10 (5): 81-86
3. Mohamad M. Alrahhal, YakoubBazi, Nassim Ammour, Mansour Zuair : Convolutional Neural Networks for Electrocardiogram Classification, 30th March, 2018, Journal of Medical and Biological Engineering,
4. Deep Convolutional Neural Network Based ECG Classification System Using Information Fusion and OneHot Encoding Techniques by Jia Li, Yujuan Si, Tao Xu,3 and Saibiao Jiang, Hindawi Mathematical Problems in Engineering, Volume 2018, Article ID 7354081, 10 pages
5. Full training convolutional neural network for ECG signals classification by KarboubKaouter, Tabaa Mohamed, DellagiSofiene, Dandache Abbas, and Moutaouakkil Fouad
6. Classification of ECG signals based on 1D convolution neural network by Dan Li ;Jianxin Zhang ; Qiang Zhang ; Xiaopeng Wei , 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)
7. Convolution Neural Network Based ECG Classifier by SharanyaS,Sridhar PA, PoornakalaJ,MuppalaVasishta,Tharani U
8. A study on user recognition using 2D ECG based on ensemble of deep convolutional neural networks by Min-Gu Kim, Hoon Ko & Sung Bum Pan, Journal of Ambient Intelligence and Humanized Computing (2020) 11:1859–1867
9. Heartbeat classification using deep residual convolutional neural network from 2-lead electrocardiogram by Zhi Li, Dengshi Zhou, Li Wan, Jian Li, WenfengMou, Journal of Electrocardiology, Volume 58, January–February 2020, Pages 105-112
10. Short Term ECG Classification with Residual-Concatenate Network and Metric Learning by Xinjing Song, Gongping Yang, Kuikui Wang, Yuwen Huang, Feng Yuan &Yilong Yin, Multimedia Tools and Applications volume 79, pages22325–22336(2020)
11. Hindawi 5198, Deep Convolutional Neural Network Based ECG Classification , Viewed 20 October 2020
12. Haykin, Neural networks. New Delhi: Pearson Education Asia, 2002.
13. P. de Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," IEEE Trans. Biomed. Eng, vol. 51, no. 7, pp. 1196–1206, Jul. 2004.
14. T. Ince, S. Kiranyaz, and M. Gabbouj, "A Generic and Robust System for Automated Patient- Specific Classification of ECG Signals," IEEE Trans. Biomed. Eng., vol. 56, no. 5, pp. 1415–1426, May 2009
15. R. M. Rangayyan, Biomedical Signal Analysis: A Case-study Approach. Wiley-Interscience, New York, pp. 18-28, 2001.
16. U. R. Acharya, P. S. Bhat, S. S. Iyengar, A. Roo, and S. Dua, "Classification of heart rate data using artificial neural network and fuzzy equivalence relation," Pattern Recognition, vol. 36, no. 1, pp. 61–68, 2003.
17. U. R. Acharya, J. S. Suri, J. A. E. Spaan, and S. M. Krishnan, Advances in Cardiac Signal Processing. Springer, 2007.
18. M. G. Khan, Rapid ECG interpretation, Third edition, 2003.
19. L. Cromwell, F. J. Weibell, and E. A. Pfeiffer, Biomedical Instrumentation and Measurements. PHI Learning Private Limited, New Delhi, 2009.
20. A. J. Moss and S. Stern, Noninvasive Electro cardiology, C. A. of Holter, Ed. London, Philadel- phia, W.B. Saunder, 1996.
21. R. S. Khadpur, Handbook of Biomedical Instrumentation. Tata McGraw Hill, 2010.
22. U. R. Acharya, J. S. Suri, J. A. E. Spaan, and S. M. Krishnan, Advances in Cardiac Signal Processing. Springer, 2007.
23. A. Cohen, Biomedical Signal Processing. CRC Press, Boca Raton, FL, 1988.
24. N. Maglaveras, T. Stamkopoulos, K. Diamantaras, C. Pappas, and M. Strintzis, "ECG pattern recognition and classification using non-linear transformations and neural networks: A review," International Journal of Medical Informatics, vol. 52, pp. 191–208, 1998.
25. S. G. Artis, R. G. Mark, and G. B. Moody, "Detection of atrial fibrillation using artificial neural networks,," Proc. Comput. Cardiology, pp. 173–176, 1991.
26. L. T. Sheffield, "Computer-Aided Electrocardiography," J Am Coll Cardiol, vol. 10, no. 2, pp. 448–455, 1987.
27. Z. Dokur and T. Olmez, "ECG beat classification by a novel hybrid neural network," Computer Methods and Programs in Biomedicine, vol. 66, pp. 167–181, 2001.
28. Z. Dokur, T. O. lmez, and E. Yazgan, "Comparison of discrete wavelet and Fourier transforms for ECG beat classification," Electron. Lett., vol. 35, no. 18, pp. 1502–1504, 1999.
29. Z. Dokur, T. Olmez, E. Yazgan, and O. Ersoy, "Detection of ECG waveforms by neural networks," Med. Eng. Phys., vol. 19, no. 8, pp. 738–741, 1997.
30. K. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tach- yarrhythmia with Fourier-transform neural network," IEEE Trans. Biomed. Eng., vol. 46, no. 2, pp. 179–185, Feb. 1999.
31. T. Olmez, Z. Dokur, and E. Yazgan, "Classification of ECG waveforms by using genetic algo- rithms," in Proc. 19th Annual Int. Conf. IEEE-EMBS, Chicago, USA, 1997, pp. 92-94.
32. B. G. Celler and P. Chazal, "Low computational cost classifiers for ECG diagnosis using neural networks," in Proc. 20th Annual Int. Conf. IEEE-EMBS, 1998, pp. 1337-1340.
33. I. Jouny, P. Hamilton, and M. Kanapathipillai, "Adaptive wavelet representation and classification of ECG signals," in Proc. 16th Annual Int. Conf. IEEE-EMBS, 1994, vol. 2, pp. 1310-1311.
34. O. Inan, L. Giovangrandi, and G. Kovacs, "Robust neural network-based classification of prema- ture ventricular contractions using wavelet transform and timing interval features," IEEE Trans. Biomed. Eng., vol. 53, no. 12, pp. 2507–2515, Dec. 2006.
35. R. Haberl, G. Jilge, and G. Steinbeck, "Spectral mapping of the electrocardiogram with Fourier transform for identification of patients with sustained ventricular tachycardia and coronary artery disease," J. Eur Heart, vol. 10, no. 4, pp. 316–322, 1989.
36. T. Olmez, "Classification of ECG waveforms by using RCE neural network and genetic algo- rithms," Electronics Letters, vol. 33, no. 18, 1997.
37. V. X. Afonso and W. J. Tompkins, "Detecting ventricular fibrillation," IEEE Eng. Med. and Biol., pp. 152–159, Mar. 1995.
38. C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," IEEE Trans. Biomed. Eng., vol. 42, no. 1, pp. 21–28, Jan. 1995.

AUTHORS PROFILE

P. Rama Santosh Naidu working as Assistant Professor in MVGR College of Engineering in the department of CSE and Part-time research scholar in Andhra University in the department of CS & SE.His research area includes Machine Learning.

Dr. G. Lavanya Devi currently working as Assistant Professor in Andhra University College of Engineering(A) in the department of CS & SE, Andhra Pradesh, India. Her research area includes Fuzzy Theory, Data Mining, Bioinformatics, Object Oriented Analysis and Design, Database Systems, Cryptography and Network Security, Algorithm Analysis, Automata Theory, Computer Communication and networks.

Dr. K.Venkata Ramana working as Assistant Professor in Andhra University College of Engineering(A), Andhra Pradesh, India. His Research area includes Data Engineering.