International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-8, June, 2019

An Implementation of Anomaly Detection in IOT DTA Using A Deep (OC-NN) With the Long Short Term Memory Network (LSTM)

K.V. Daya Sagar, DBK Kamesh

Abstract: An Electrocardiography (ECG) signals are accessed mainly to monitor the health condition of the human heart, and the resulting time series signals are analyzed manually by the medical professionals to detect if there are any anomalies such as arrhythmia. Manual Diagnosis of ECG Signals has been often Prone to Errors. An Electrocardiography (ECG) signals are accessed mainly to monitor the health condition of the human heart, and the resulting time series signals are analyzed manually by the medical professionals to detect if there are any anomalies such as arrhythmia. Manual Diagnosis of ECG Signals has been often Prone to Errors. Past work in Automating the Analysis requires extensive Pre-Processing, which is time-consuming and cumbersome. It takes significant time for Heart Patients in their Precarious Condition. There is a requirement of Computational Analysis, which is fast and efficient. Some of the analysis develops the marked features and design a classifier for discriminating between the healthy ECG signals and those which contains Arrhythmia. This method requires knowledge and relevant data of the various types of Arrhythmia of training the model.However, there can be many different and different new types of Arrhythmia can occur, which previously were not a part of the original training set. Thus, it may be wiser to adopt an anomaly detection approach to analyzing them. In this paper, we are utilizing A deep one class neural network (OC-NN) architecture with the Long Short-Term Memory Network (LSTM) units for developing a predictive model from the healthy ECG signals. The probability distribution of the prediction errors from the models, using the Maximum Likelihood Estimate (MLE) is used for indicating anomalous or non-anomalous behavior. The main advantage of using LSTM networks is that the ECG signals be directly applied to the system without any extensive pre-processing as used by other Detection techniques. No Prior information of abnormal signals makes it worthwhile, as it needs to be trained only on average datza. MIT-BIH Arrhythmia Physionet Database has been used to obtain ECG time series data for both non-anomalous periods and irregular periods. Both the Stateful and homeless Modes of LSTM are projected and enforced. The Results from the Stateful LSTM Model show Precision of 99.36% and a TPR/FPR ratio of 145.39. Results are promising and indicate that the Deep-Stacked Long Short-Term Memory Networks (LSTM) models are feasible for detecting anomalies in ECG signals within a short period.

Index Terms: one class SVM, anomalies detection, outlier detection, deep learning, LSTM, OC-NN.

I. INTRODUCTION

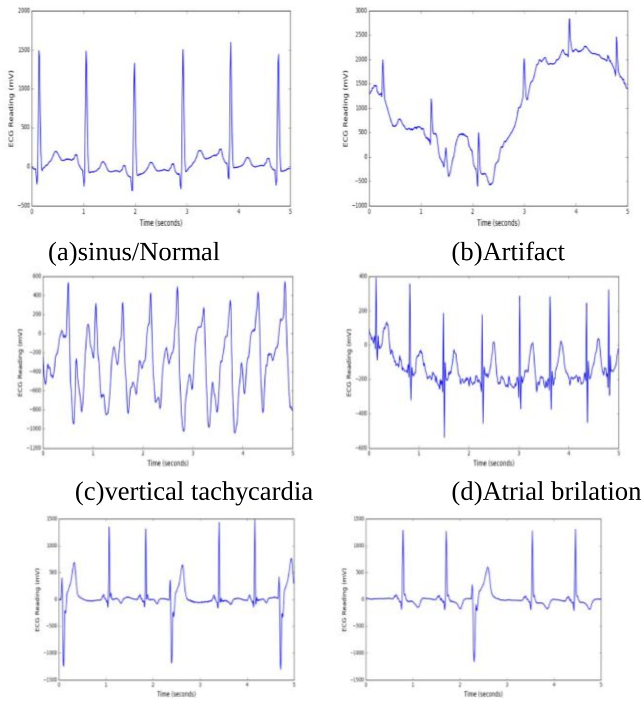
Most of the time, IoT data exceptionally patient monitoring system a frequent need to analyze if any outlier or error is determined. Such situations are known as anomalies.

Revised Manuscript Received on June 05, 2019

Dr. Duvvuri b k kamesh, Professor, Mall Reddy Engineering College for Women, Maisammaguda, Secunderabad, he has the research guide of Shri Venkateshwara University, Gajraula, Uttar Pradesh ,India,

K.V.Daya Sagar, Research Scholar, Shri Venkateshwara University, Gajraula, Uttar Pradesh., India,

The goal of anomaly detection is to identify all cases in a data-driven fashion Chandola et al. [2007]. Anomalies can be caused by errors in the data but sometimes are indicative of new, previously unknown, in fact, Hawkins [1980] characterizes an anomaly as a perception that veers off so remarkably from different contemplations as to stimulate doubt that an alternate component created it. The rhythm of a human heart is regulated by electrical signals produced by two nodes within the heart and conducted through a series of specialized cardiac cells. During healthy, normal operation, this occurs at regular intervals and the electrical signal, which causes the heart muscles to contract, propagates via the cardiac electrical conduction system along the correct path through the atria and ventricles. Figure 1.1 shows an example of rhythm types [1].



(a)sinus/Normal (b)Artifact

(c)vertical tachycardia (d)Atrial brilation

(e)ventricular bigeminy (f)Premature ventricular contraction

Figure :1.1 Examples of rhythm types

One-class Support Vector Machines (OC-SVM) Scholkopf¨ and Smola [2002], Tax and Duin [2004] ar wide used, effective unsupervised techniques to spot anomalies. However, the performance of OC-SVM is sub-optimal on advanced, high dimensional datasets Vapnik, and Vapnik [1998], Vishwanathan et al. [2003], Bengio et al. [2007]. Leveraging multi-channel time series data in clinical settings for various classification tasks is an active area of research [10].

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

In recent work, this approach has been used to classify human activities based on motion sensor data [11]. Existing techniques for ECG classification in particular almost exclusively use single-channel data [5{7], however, likely due to lack of data. The ECG recordings provided by the PhysioNet database of physiological signals [12] Since the hybrid models extract in-depth options victimization associate autoencoder so feed it to a separate anomaly detection methodology like OC-SVM, they fail to influence naturalistic learning within the hidden layers. During this paper, we tend to devolve on the speculation to integrate associate OC-SVM equivalent objective into the neural specification. The OC-NN combines the flexibility of deep networks to extract a more and more wealthy illustration of information together with the one-class aim, that obtains the hyperplane to separate all the typical information points from the origin. The OC-NN approach is novel for the following crucial reason: information illustration is driven by the OC-NN objective and is therefore tailor-made for anomaly detection. We tend to show that OC-NN can do a comparable or higher performance in some eventualities than existing shallow progressive ways for Io Tbased mostly EKG information set whereas having consistent coaching and testing time compared to the prevailing ways.

The rest of the paper is organized as follows. In Section 2, we presented the complete survey of the anomaly detection model, and the main OC-NN with LSTM combined model is developed in Section 3. The experiment setup, evaluation metrics, and model configurations are described in Section 4. The results and analysis of the experiments are the focus of Section 5. We conclude in Section 6 with a summary and directions for future work.

II. RELATED WORK

As we have access to multi-channel data, we incorporate this increased dimensionality into our algorithm, in contrast to the single-channel input format commonly used in ECG classification. Individual channels record cardiac electrical activity from various spatial angles, and the use of multiple channels is likely to give a more in-depth insight into any underlying patterns of arrhythmia that can be interpreted by our model. We train two baseline models: a single-layer, unidirectional LSTM, and a convolutional neural network without residual connections. We then train models from three categories (LSTM only, residual networks, and LSTM-OCNN), for a total of six models, not including our baseline models. All models exceed baseline performance on training, validation, and test accuracy. A 2-layer, bidirectional LSTM achieves the best performance overall; through networks with residual structures achieve higher training accuracy, their validation and test accuracies are Lower than those of the LSTM only systems. The overall F1 score of 0.804 achieved by this network exceeds reported cardiologist F1 scores reported in [2].

In general, much work in this area, until recently, has relied on extracting hand-crafted features and comprehensive prior knowledge of specific arrhythmias and their waveform patterns. This limits the number of classes that can be reliably differentiated with a given model, as the morphology of different waveforms is usually highly specific to a given arrhythmia. More recent developments in arrhythmia classification research have utilized well-known deep learning algorithms such as deep OCNN's and incorporation of skip/residual layers to improve classification accuracy [2]. Existing techniques for ECG classification almost exclusively use single-channel data [5,7], however, likely due to lack of data. The ECG recordings provided by the PhysioNet database of physiological signals [12], used for testing and training many ECG classification algorithms, are primarily dual or single channel.

III. METHODOLOGY

The data used to train the models is a set of ECG recordings, obtained from IoT based cloud patient monitoring system, for 106 patients for varying lengths of time. Seven channels are provided, though one is removed during extraction as it is merely a linear combination of three of the other channel readings. These include manually labeled alarms of specified length this is distinct from many of the datasets found in the MIT-BIH database [12] that is typically used for ECG classification tasks, as the MIT-BIH data usually does not provide duration information for any of the alarms. The window for each sample is either truncated or padded to 5 seconds as many of the viable `alarm segments '|namely, signals not caused by internal errors are close to this length. As this data was passively collected, the distribution of classes is extremely skewed, so more common types (accurately, sinus rhythm and noise) were downsampled to achieve a more even class distribution. We extract over 151,000 segments classified as one of six classes; from this dataset, we sample nearly 13,000 data points due to the class above imbalance. We split this dataset into training, validation, and test sets, at 69%, 16%, and 14% respectively (obtained from an initial 86/14 divided into practice and test data, followed by an 80/20 split of that training data into training and validation sets). To achieve reasonable training time, the samples were decimated by a factor of 2 to make an effective sampling rate of 120 Hz from the original 240 Hz. Each sample, thus, is of size (606). Each sample is then assigned a ground truth label from its corresponding alarm, regardless of the length of the actual signal (as the data is padded if it is shorter than our 5-second window). Here table 3.1 shows the exact profile of the ECG data set.

Type	Original	Downsampled
Sinus rhythm	126,435	4,000
Artifact/Noise	18,953	3,000
Ventricular tachycardia	249	249
Atrial fibrillation	1,452	1,452
Bigeminy	896	896
PVC	2,374	2,374
Total	150,359	11,971

Table 3.1 ECG data set profile

III.1. MODEL ARCHITECTURES

III.1.1.FROM ONE CLASS SVM TO ONE CLASS NEURAL NETWORKS

We now present our one-class Neural Network (OC-NN) model for unsupervised anomaly detection.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The method is designing a neural architecture using an OC-SVM equivalent loss function. Using OC-NN, we will be able to exploit and refine features obtained from unsupervised transfer learning specifically for anomaly detection. This, in turn, will it make it possible to discern anomalies in complex data sets where the decision boundary between normal and abnormal is highly nonlinear.

IIII.1.1.2.LSTM

Long short-term memory units are used to build one class neural networks. They address the issue of exploding/vanishing gradients in standard network architectures. Several gates control updates from inputs, as well as how much of a memory state is retained and passed on to the proceeding timestep. Notably, the cell state C_t (for timestep *t*) is only ever gated elementwise with the forget gate *f* and the input activation gate *g* without undergoing a linear transformation. This allows gradients to pass through many steps in time without vanishing (hence `memory'), and thus, the model can learn long-term dependencies within the data. Formally, we define the relationship between cell inputs x*t*, h_{t-1}, $c_{t-1,}$ and cell outputs h_t, c_t for data inputs x_t, hidden states h_t, and cell states c_t at a given timestep *t* as follows

$$
\begin{aligned}
i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) = \sigma(\hat{i}_t) \\
f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f) = \sigma(\hat{f}_t) \\
o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o) = \sigma(\hat{o}_t) \\
g_t &= \tanh(W_g x_t + U_g h_{t-1} + b_g) = \tanh(\hat{g}_t)
\end{aligned} \quad (1)
$$

$$
\begin{aligned}
c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\
h_t &= o_t \odot \tanh(c_t)
\end{aligned} \quad (2)
$$

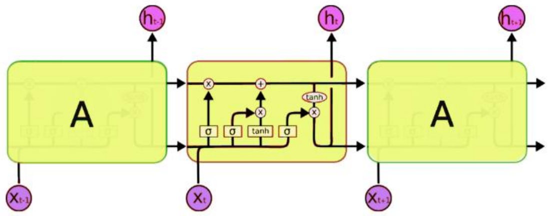


Figure 3.1.2 The architecture of LSTM network

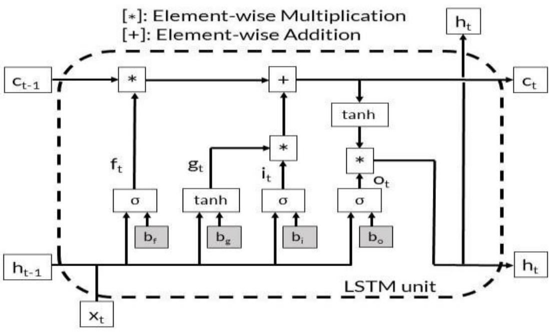


Figure 3.1.3. LSTM Unit

Long short term memory units are used to build recurrent neural networks. They address the issue of exploding/vanishing gradients in standard recurrent architectures. Several gates control updates from inputs, as well as how much of a memory state is retained and passed on to the proceeding timestep. For our LSTM networks, an individual sample is treated as a 600-timestep sequence, where each timestep is a 6-feature data point. In addition to a single-layer unidirectional baseline model, we also train 2- and 5-layer stacks of LSTM units, where the output of each LSTM unit not in the final layer is treated as input to a group in the next. To increase the contextual information available to the model, we also train a bidirectional variant of the 2-layer LSTM model. These use the same update equations but add a second LSTM owing backward in time; cell outputs from both networks are then concatenated at each time step.

III.3.1.3. OC-NN Algorithm

We summarize the solution in Algorithm 1. We initializer(0) in Line 2. We learn the parameters(*w*, V) of the neural network using the standard Backpropagation(BP) algorithm (Line 7). In the experiment section, we will train the model using features extracted from an autoencoder instead of raw data points. However, this has no impact on OC-algorithm. As shown in Theorem 3.1, we solve for r using the -quantile of the scores in. Once the convergence criterion is satisfied, the data points are labeled normal or abnormal using the decision function S*n* = sgn(^y_n -r).

ALGORITHM 1: ALGORITHM FOR ONE-CLASS NEURAL NETWORK (OC-NN)

1. Input: Set of points X_n:, n : 1, …, N
2. Output: A Set of decision scores S_n: = $_y$^$_n$:, n: 1,...,N for X
3. Initialise $r^{(0)}$
4. *t* ← 0
5. while (no convergence achieved) do
6. Find ($w^{(t+1)}$, $V^{(t+1))}$
7. r^{t+1}←v^{th} quantile of {y^n^{t+1}gn}N *n*=1
8. t← t + 1
9. end
10. Compute decision score S_n: = y^$_n$ - r for each X_n:
11. if (S_n: ≥ 0) then
12. X_n: is the normal point
13. else
14. X_n: is anomalous
15. return {S_n}

IV. EXPERIMENTAL SETUP

In this section, we show the actual experimental setup required for the patients to monitor ECG signals and collect the ECG data through the AD8232 ECG module and stored into the cloud as data set. In future, anybody wants to do some analysis against that data they can do and apply OC-NN formulation over the state-of-the-art methods on real-world data. Our proposed model is compared with various state of the art methods given bellow.

Figure 4.1 Complete Hardware Kit with connecting sensors

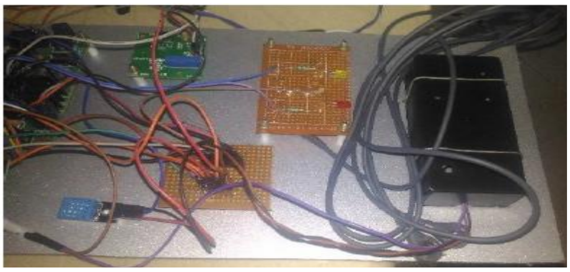


Figure 4.2 AD8232 ECG Module

IV.1 METHODS COMPARED

The available deep learning models are compared with each other and determine the instances and nature of each model and find the anomaly percentage and attribute of trained data sets

Dataset	# instances	# anomalies	# features
OC-NN	190	10	512
SVM	single class	1% (from all class)	784
LSTM	single class	10% (from all class)	3072
OC-NN LSTM	1050 (stop signs)	100 (boundary attack)	3072

Table 4.1: Summary of datasets used in experiments.

We formulate this as a sequence classification problem and train multiple models to investigate the effectiveness of a combining OC-NN connections with LSTM architecture. We train six model architectures from three categories:

1.LSTM

2-layer unidirectional LSTM
5-layer unidirectional LSTM
2-layer bidirectional LSTM

2.OC-NN

16-block deep one class neural networks

3.**Combined Models**

Unidirectional LSTM-OCNN
Bidirectional LSTM – OCNN

IV.1.2. PARAMETERS OF BASELINE MODEL

The proposed OC-NN method is compared with several state-of-the-art baseline models, as illustrated in Table 4.1. The model parameters of shallow baseline methods are used as per implementation in Ruff et al. [2018]. Shallow Baselines As baselines, we train a 4-block (each block containing a sequence of convolution, max-pool, batch normalization, and dropout layers) with maxpool OC-NN connections as well as a single-layer, unidirectional LSTM with a hidden dimension of 100.

Table 4.1 Aggregate test set accuracy for all classes using multi-channel data

IV.1.3.TRAINING

All networks are trained using a minibatch size of 64, and 100 epochs of training. We minimize categorical cross-entropy loss in all cases, and we use the Adam optimizer [20], an extension of the classic stochastic gradient descent optimizer. All LSTM networks use a hidden dimension of 100, as well as a small dropout parameter (p = 0:1) for all gates (and thus applied between cell input and output as well as between `unrolled' instances of each unit). All dropout layers in OC-NN models used p = 0:5 to maximize regularization [21]. To investigate the usefulness of multi-channel ECG data, we also retrain a representative model from each architecture using only one of the available channels per sample (specifically, the BDLSTM with two stacked LSTM layers, deep one class neural network, and combined LSTM-OCNN with two stacked LSTM layers).

V. EXPERIMENTAL RESULTS

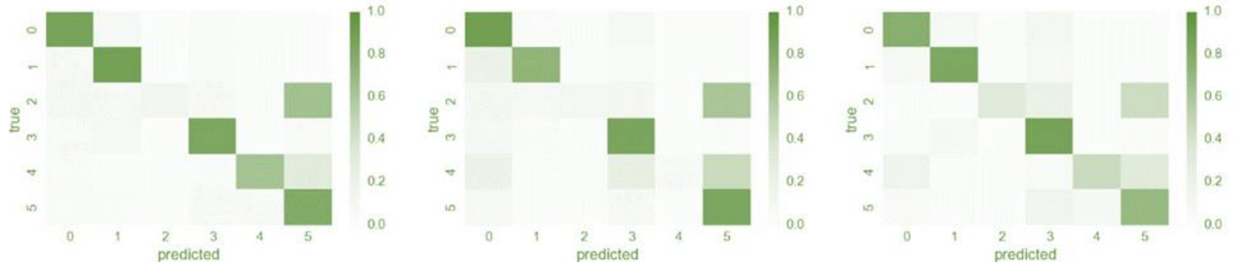
In this section, we present empirical results produced by OC-NN model on real-time data. The tables of this section summarize our experimental results. We investigate the overall accuracy of all models (table 4.1) and derive more detailed performance metrics (recall, precision, septicity, and F1 score) of the representative models mentioned above (table 5.3). A comparison between F1 scores of the same three models using both multi-channel (6 channel) and single-channel data is also provided (table 5.4).

Except for VT (whose classification showed relatively poor performance in all models) and PVC, the 2-layer bidirectional LSTM had the highest F1 score out of the three representative models, as shown in table 5.2. The low F1 scores for VT classification are likely a result of its low support in the dataset. This is also seen to a lesser extent in bigeminy classification, with low recall and F1 scores in the one class neural network. From Figure 4.1, we see that VT is often mis labeled as PVC; both these arrhythmias have wide QRS complexes [3], which are the characteristic waveform morphologies visible.

F1 Score Class Comparison			
Rhythm class	BDLSTM	OC-NN	LSTM-OC NN
Sinus rhythm	0.832	0.754	0.783
Artifact/Noise	0.854	0.808	0.823
Ventricular tachycardia	0.225	0.069	0.407
Atrial brillation	0.827	0.783	0.774
Bigeminy	0.683	0.116	0.543
PVC	0.779	0.801	0.714
Overall	0.803	0.718	0.752

Table 5.1 F1 score comparison over classes for the test set, using representative LSTM (2-layer BDLSTM), residual (OC-NN), and combined (LSTM-OC NN) Models. Overall F1 score is a weighted average given the class's support in the test set.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication



(a)BDLSTM (b)OC-NN (c)LSTM-OCNN

Figure 5.1: Confusion matrices for the three representative models where each row (true class) is scaled by its sum, for visualization of classification recall. True and predicted labels are numbered in the order presented in tables 5.1-5.3.
Figure 5.1: Confusion matrices for the three representative models where each row (true class) is scaled by its sum, for visualization of classification recall. True and predicted labels are numbered in the order presented in tables 5.1-5.3.

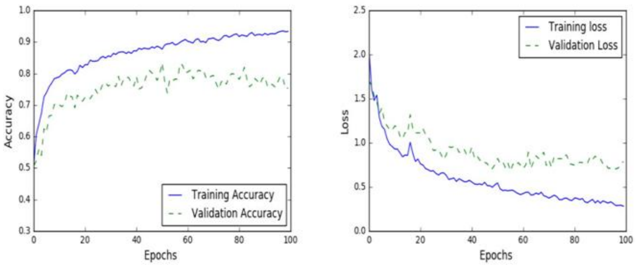
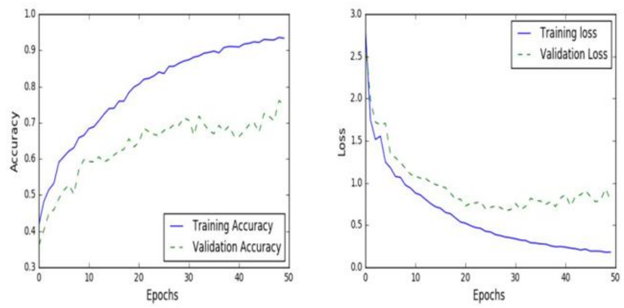


Figure 5.2: Deep residual OC-NN accuracy
and loss curves: This model also shows a degree of over thing as validation accuracy does not show a considerable increase after about 40-50 epochs, even as performance on training data continued to improve.

We note that the overall F1 score of 0.803 achieved by the 2-layer BDLSTM exceeds cardiologist scores as reported in [2], both sequence level1 F1 score of 0.719 and set level2 F1 score of 0.751.



(a)Accuracy Curves (b)Loss Curves

Figure 5.3: Combined BDLSTM-OCNN accuracy
and loss curves: Much like the deep residual network, this model also shows a degree of over tting as performance on validation data does not improve after about 30 epochs. Validation loss started to increase slightly as well.

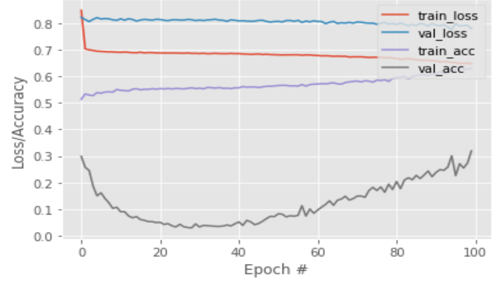


Figure 5.4: shows Fake noise LSTM_OCNN training loss and accuracy

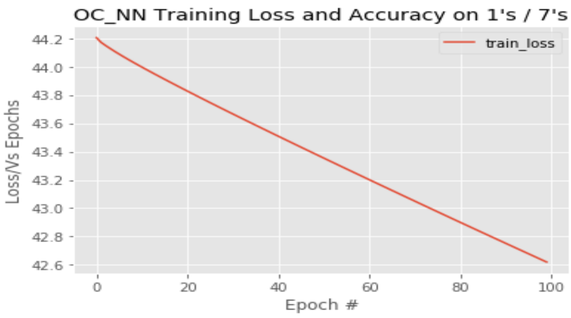


Figure 5.5 shows OC-NN Training Loss and accuracy

VI. CONCLUSION AND FUTURE WORK

We show that using known deep neural network algorithms for classifying time series data like IoT sensors generated ECG signals allows for accurate classification of normal, benign, and critical arrhythmias as well as distinguishing artifacts and noise from multi-channel ECG recordings. A layered bidirectional LSTM, as well as a combined LSTM-OCNN architecture, is able to achieve relatively high accuracy and precision without the use of feature engineering or extraction of previously known waveform patterns. We also show that ECG classification greatly benefits from the use of multi-channel data, with nearly all classes and models showing markedly decreased accuracy when only one channel is used.

REFERENCES

1. Varun Chandola, Arindam Banerjee, and Vipin Kumar. Outlier detection: A survey. ACM Computing Surveys, 2007.
2. Rajpurkar, P., Hannun, A. Y., Haghpanahi, M., Bourn, C. & Ng, A. Y. Cardiologist-level arrhythmia detection with convolutional neural networks (2017).
3. Douglas M Hawkins. Identification of outliers, volume 11. Springer, 1980.
4. Bernhard Scholkopf¨ and Alexander J Smola. Support vector machines, regularization, optimization, and beyond. MIT Press, 656:657, 2002.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

5. David MJ Tax and Robert PW Duin. Support vector data description. Machine learning, 54(1):45–66, 2004.
6. Bortolan, G., Degani, R. & Willems, J. L. Neural networks for ECG classi_cation in Computers in Cardiology 1990, Proceedings. (1990), 269-272.
7. Vladimir Naumovich Vapnik and Vladimir Vapnik. Statistical learning theory, volume 1. Wiley New York, 1998.
8. Cheng, W. & Chan, K. Classi_cation of electrocardiogram using hidden Markov models in Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual
9. International Conference of the IEEE (1998), 143{146.
10. SVN Vishwanathan, Alexander J Smola, and M Narasimha Murty. Simplesvm. In Proceedings of the Twentieth International Conference on International Conference on Machine Learning, pages 760–767. AAAI press, 2003.
11. Goldberger, A. L. et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101, e215(e220(2000 (June 13)).
12. Yoshua Bengio, Yann LeCun, et al. Scaling learning algorithms towards ai. Large-scale kernel machines, 34(5):1–41, 2007.
13. Chong Zhou and Randy C Paffenroth. Anomaly detection with robust deep autoencoders. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 665–674. ACM, 2017.
14. Raghavendra Chalapathy, Aditya Krishna Menon, and Sanjay Chawla. Robust, deep, and inductive anomaly detection.arXiv preprint arXiv:1704.06743, 2017.
15. Jerone TA Andrews, Edward J Morton, and Lewis D Griffin. Detecting anomalous data using auto-encoders. International Journal of Machine Learning and Computing, 6(1):21, 2016a.
16. Sarah M Erfani, Sutharshan Rajasegarar, Shanika Karunasekera, and Christopher Leckie. High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. Pattern Recognition, 58:121–134, 2016.
17. Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10):1345–1359, 2010.
18. Jerone TA Andrews, Thomas Tanay, Edward J Morton, and Lewis D Griffin. Transfer representation-learning for anomaly detection. ICML, 2016b.
19. Charu C Aggarwal. Outlier Analysis. Springer, 2nd edition, 2016.
20. Bernhard Scholkopf,¨ John C Platt, John C Shawe-Taylor, Alex J Smola, and Robert C Williamson. Estimating the support of a high-dimensional distribution. Neural Computation, 13(7):1443–1471, Jul 2001.
21. LJ Cao, Kok Seng Chua, WK Chong, HP Lee, and QM Gu. A comparison of PCA, PCA, and ica for dimensionality reduction in the support vector machine. Neurocomputing, 55(1-2):321–336, 2003.
22. Julia Neumann, Christoph Schnorr,¨ and Gabriele Steidl. Combined SVM-based feature selection and classification. Machine learning, 61(1-3):129–150, 2005.
23. Yong Shean Chong and Yong Haur Tay. Abnormal event detection
24. in videos using spatiotemporal autoencoder. In International Symposium on Neural Networks, pages 189–196. Springer, 2017.
25. Erik Marchi, Fabio Vesperini, Stefano Squartini, and Bjorn¨ Schuller. Deep recurrent neural network-based autoen-coders for acoustic novelty detection. Computational intelligence and neuroscience, 2017, 2017.
26. Muhammad Sohaib, Cheol-Hong Kim, and Jong-Myon Kim. A hybrid feature model and deep-learning-based bearing fault diagnosis. Sensors, 17(12):2876, 2017.
27. Lukas Ruff, Robert Vandermeulen, Nico Goernitz, Lucas Deecke, Shoaib Ahmed Siddiqui, Alexander Binder, Emmanuel Muller,¨ and Marius Kloft. Deep one-class classification. In Jennifer Dy and Andreas Krause, editors, Proceedings of the 35th International Conference on Machine Learning, volume 80 of Proceedings of Machine Learning Research, pages 4393–4402, Stockholmsmassan,¨ Stockholm Sweden, 10–15 Jul 2018. PMLR. URL http://proceedings.mlr.press/v80/ruff18a.html.
28. Huan Xu, Constantine Caramanis, and Sujay Sanghavi. Robust PCA via outlier pursuit. In Advances in Neural Information Processing Systems, pages 2496–2504, 2010.
29. Stephen Boyd and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004.
30. Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation forest. In Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on, pages 413–422. IEEE, 2008.
31. Emanuel Parzen. On the estimation of a probability density function and mode. The annals of mathematical statistics, 33(3):1065–1076, 1962.
32. Jonathan Masci, Ueli Meier, Dan Cires¸an, and Jurgen¨ Schmidhuber. Stacked convolutional auto-encoders for hierarchical feature extraction. In International Conference on Artificial Neural Networks, pages 52–59. Springer, 2011.
33. Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015.
34. Franc¸ois Chollet et al. Keras. https://github.com/keras-team/keras, 2015.
35. Mart´ın Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467, 2016.
36. Yann LeCun, Corinna Cortes, and Christopher JC Burges. Mnist handwritten digit database. AT&T Labs [Online].Available: http://yann. lecun. com/exdb/mnist, 2, 2010.
37. Johannes Stallkamp, Marc Schlipsing, Jan Salmen, and Christian Igel. The German traffic sign recognition benchmark: a multi-class classification competition. In Neural Networks (IJCNN), The 2011 International Joint Conference on, pages 1453–1460. IEEE, 2011.

AUTHORS PROFILE

K.V.Daya Sagar, Research Scholar, Shri Venkateshwara University, Gajraula, Uttar Pradesh., India, his research interests are Cloud Computing, IoT, Cyber Security, Machine Learning, and deep learning.

Dr. DUVVURI B K KAMESH, Professor, Mall Reddy Engineering College for Women, Maisammaguda, Secunderabad, he has the research guide of Shri Venkateshwara University, Gajraula, Uttar Pradesh., India, his research interests are Cloud Computing, IoT, Cyber Security, Machine Learning and deep learning, embedded systems and software engineering.