Accuracy Enhancement during Atrial Fibrillation Detection using Hybrid Machine Learning Algorithm and Echo Peak Detection Algorithm

Alugonda Rajani
Assistant Professor, ECE Department
JNTUK
Kakinada, India

Sai Lakshmi Dangeti
Student, ECE Department
JNTUK
Kakinada, India

Abstract—**The medical instrument Electrocardiogram (ECG) shows the heart functioning. Monitoring and analysis of ECG signal is important for diagnosing the cardiac disorders like Atrial fibrillation (AFib). AFib is a quivering or irregular heartbeat that can lead to blood clots, stroke, heart-failure and other heart related complications if gone untreated. Poor signal quality is the chief limitation of efficacy of biological signal analysis. The accuracy of AFib detection is to be increased by detecting and reducing the false alarms during patient monitoring. The peak points from ECG are detected using the Pan Tompkins Algorithm. It gives amplitude and location information. AFib Detection Algorithm is necessary for statistical measures of the peak points. Threshold values are used for atrial fibrillation detection. Deep belief networks (DBN) are the classifiers used to detect accuracy of the readings. An accuracy of 85% is possible by making use of a hybrid DBN. The software used for implementation is MATLAB. Application of Potential Detection Algorithm will detect echo peak present in the signal and amplified accuracy can be obtained. The raise in accuracy through Atrial Fibrillation detection ensure patient protection and preventfright for monitoring staff.**

Keywords—Electrocardiogram(ECG), Atrial Fibrillation(AFib), Echo Peak Detection, Hybrid DBN, Accuracy

INTRODUCTION

Atrial Fibrillation (AFib) is an anomalous heart pace categorized by brisk and asymmetrical heartbeats. The medical equipment ECG is the most frequently used tests by the healthcare professionals for evaluation of patient condition and it is also a vital part of cardiac appraisal in modern era. Willem Einthoven is the founder and father of modern ECG. Through this device, presence of AFib can be predicted visually.

The rapid and irregular activity of the heart during Afib increases the variability and complexity of RR interval series [5]. So these both must be of major consideration during Afib detection. Afib generally proves to be less critical for healthy people rather than those suffering with other major health ailments related to heart.

If not treated in time, it might prove to be hazardous and can cause loss of life at times. So an early and accurate detection of AFib is very essential for the patient.

The normal ECG generally consist of P,Q,R,S,T points as depicted in Fig 1. The P wave represents the depolarization of the left and right atrium furthermore also corresponds to atrial retrenchment. The repolarization of left and right atrium is so minute that it can't be made visible on ECG.

The QRS complex represent the electric impulse as it spreads through the ventricles and indicates ventricular depolarization. The T point in the wave comes after thethe QRS composite and implies ventricular repolarization.AFib has a irregular rhythm with an absence of P wave and this is replaced by random waves unlike Normal Sinus Rhythm (NSR) and its pictorial representation is given in Fig 2.

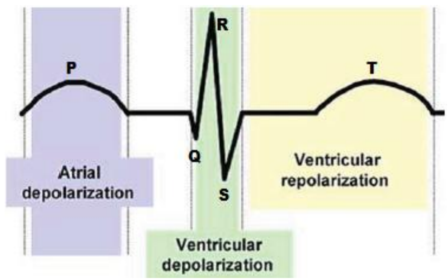


Fig 1. Standard Form of ECG

Atrial fibrillation (AF) is the largely occurred clinical arrhythmia distressingaround 3 million Americans. It have a predominance of 18 % and an occurrenceof 21/1000 patient years in thoseadults aging than 85. At an age of 55, the lifetime menace of budding AF is roughly 23% [21].

Around2% of individuals younger than 65 are suffering with AFib and around 9% of populace aged 65 years or older than 65 [1]. Thepremature discovery of AFib is vital for those who endure cardiovascular diseases, the aged or stroke patients towhom second stroke anticipation is of principal significance [2].

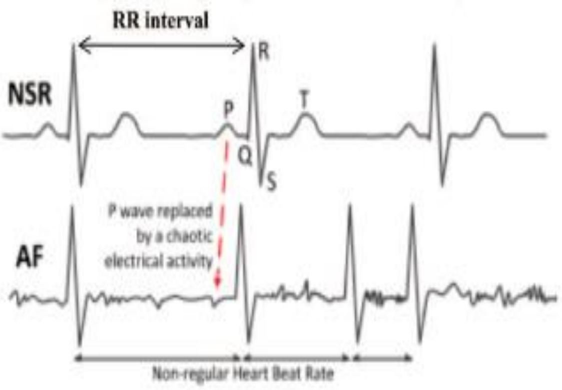


Fig 2. Comparison of NSR and AFib

Atrial fibrillation is an independent risk factor for death (relative risk in men 1.5 and in women 1.9) [20] and a major cause of ischemic stroke whose impact increases with age, reaching 23.5% in patients older than 80 years [22]. Accurate detection of AF is crucial since treatment options such as chronic anticoagulation, ant arrhythmic therapy and radiofrequency ablation offer significant benefits but also carry potentially serious risks.

Multiple algorithms for automatic detection of AFib have been developed and tested in terms of various statistical measures of performance. But machine learning algorithms have become the quickest growing fields in engineering due to their robust nature and simplified designs in accomplishing a complex task. Convergence is one of the major concepts for contemporary technology [8]. So a hybrid form of machine learning is implemented in this paper to enhance the accuracy of AFib detection with the help of a DBN classifier.

The remaining sections depicts the background work for machine learning, methodology that has been implemented, experimental results, discussion regarding the results and finally the conclusions in sections II-V respectively

ATRIAL FIBRILLATION DETECTION ALGORITHM

There are many prevailing methods for atrial fibrillation detection briefed out in [5],[15], [16], [17].

This paper made use of the statistical method discussed by Dash in [17].The three statistical tests used were: TurningPoints Ratio (TPR); a nonparametric test which characterizesrandomness of the RR intervals series, Root Mean Squares ofSuccessive Differences (RMSSD); a parametric test whichquantifies its variability, and Shannon Entropy (SE); aparametric test which characterizes its complexity.The key downside of all the algorithms except [17] is that they are reliant on the heftiness of the training data. For instance, if the features of AF are altered from those which are learned in the training data, the accuracy of AF detection is conciliated. All the below listed parameters can be calculated from formulae represented in [17].

RMSSD:-

$$\sqrt{\frac{1}{l-1}\sum_{i=1}^{l-1}\left(a(i+1)-a(i)\right)^{2}}(1)$$

TPR:-

$$\frac{2l-4}{3}\pm\frac{16l-29}{90}(2)$$

SE:-

$$\sum_{i=1}^{B}p(i)\,\frac{\log(p(i))}{\log\frac{1}{B}}(3)$$

where p(i), probability distribution is given by

$$\frac{N\text{bin}(i)}{l-N\text{outliers}}(4)$$

Here,
l=segment length in beats =128 [17].
a(i) is the time at which R peak arrives
N_{bin} is the number of heart beats in the i_{th} bin
$N_{outliers}$ is the number of outliers= 16 [17].
B is thenumber of bins= 16 [17].

PARAMETERS WHICH ARE NEEDED TO CLASSIFY AN L-BEAT RR INTERVAL SEGMENT OF ECG AS AF.

Parameter	***Meaning***	***Range of values for declaring a segment as having AFib***
RMSSD	Root Mean Square of Successive Differences	RMSSD>0.1*meanRR
TPR	Turning point ratio	0.54<TPR<0.77
SE	Shannon entropy	SE>0.7
l	Segment length in beats	128

In the contemporary cram, we utilize a blend of three diverse numerical methods proficient of exposing the incidence of arbitrariness in a signal. By means of such an approach, there will be some curtail in the need for widespread storage capability while preserving the accuracy of the detection. A beat-by-beat scrutiny of the detection results is offered and accuracy is made known during a ROC curve examination. In addition we also used an ectopic beat filtering scheme to prevent misdetection of ectopic rhythms as AF. The ROC analysis revealed that the optimal segment length is 128 RR intervals with at least 50% AF to ensure correct classification of the segment as AF.

METHODOLOGY

Database

We have used MIT-BIH AFDB available in PhysioNet [3] in our algorithm. This database consists of 25 Holter ECG files of human subject having AFib. Every recording is around 10 h in duration, with a sampling frequency of value 250 Hz and 12-b resolution above a range of ±10 mV.

All recordings are completely annotated and may perhaps contain normal sinus rhythm, AFib, and atrial flutter.

Pre-treatment of Signal

The available AFib database is never a clean signal and in addition it contains baseline wander (BW), muscle noise, power line interference and patient wandering disturbances. It would be difficult for diagnosis of signal with such faults. So the signal undergoes filtering procedure so as to remove all the unnecessary noise in the signal.Muscle noise and power

line interferences are the high frequency noise which can be reduced using low pass filter of frequency 11 Hz for an estimate. Baseline wander can be eliminated using the smoothing filter. In this work, a bandpass filter is used which has both the properties of high pass filter and low pass filter.
This filter is designed based on the noise filtering step of Pan-Tompkin QRS detection algorithm [18]. To use AFib detection algorithm, identifying R peaks is necessary. The R peaks are identified using the Complete Pan-TompkinImplementation ECG QRS detector toolbox [18], [19]. This consists of bandpass filter which was discussed earlier.mAn example of ECG signal filtering using low pass and high pass filter is shown below is shown here in Fig 3.

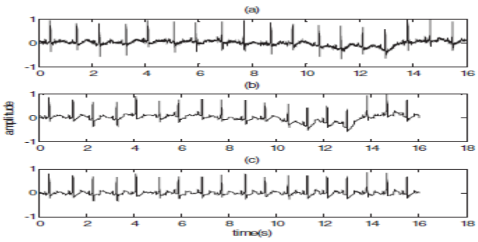


Fig 3. (a). Original ECG. (b). Low Pass Filtered. (c). High Pass Filtered

Signal Contamination
An algorithm is generally required to detect, identify, quantify and mitigate motion artifacts in the signal. For this a true signal is contaminated with previously recorded motion artifact or simulated motion artifact. Motion artifact contamination adversely affects the biological signal interpretation [10]. Motion artifacts generally lead to misinterpretation of patient condition which in turn increases strain for health care professionals.
Filtering the signals alone is not sufficient because the motion artifacts caused due to movement of electrodes with respect to skin and movement of electrode cables will partly cover the ECG in both time and frequency domain. Replicated signals give better appraisal. Simulations of motion artifact contaminated data can be obtained by adding motion artifact noise to a clean ECG recording which is here obtained after the filtering process done earlier.
An electrode motion artifact signal is generated using the autoregressive (AR) model of the noise recording. It usually resembles a random process where the output will belinearly dependent on the input values. The below given equation shows the AR model for an order value of p.

$$X(n) = \sum_{i=1}^{p} \varphi i\, X(n-i) + \varepsilon(n)$$
(5)

Where X(n) is the obtained contaminated signal$\varphi_i$indicates a set of AR coefficients which are obtained using the Yule Walker's method and ε(n) represents a random zero mean white noise. In this paper the order of AR model is considered to be 20 [23] by making use of Akaike's Information Criterion (AIC). This evaluates the quality of a simulated model for a given data.
The contaminated signal is obtained and this again undergoes filtering process for echo peak detection using the echo peak detection algorithm so that all the P peaks and maximum peaks generated due to the echoes generated which are indirectly considered a noise, unwanted signal so as to remove them later by defining threshold values. are also properly enhanced so that any false alarm that arise can be gated easily using the DBN architecture which is discussed below.

Gating of Signal containing no AFib in the segments
The AFib detection algorithm explainer prior to this is applied to the so obtained contaminated signal. A threshold is set up for the purpose of gating.
A DBN [4] is trained using this threshold value for gating purpose. Ten fold cross validation is performed and also contrastive divergence method [24] is made use of for the purpose of training RBM's through a 200 epoch attempt. Fine tuning has demanded for a back propagation method [25] here. The implemented methodology in the form of a block diagram is presented in Fig 4.Band pass filter is the filter which is used in addition to SavitzkyGolay filter for the purpose of enhancing peaks in the corresponding ECG segments.
The DBN classification of the corresponding segments of ECG is made use of for the purpose of gating the AFib alarms arising that were due to segments having no AFib.The basics of machine learning can be studied in

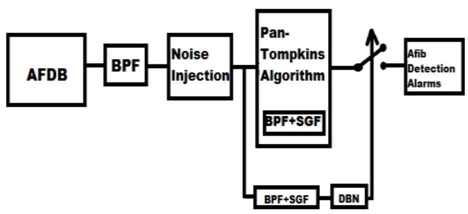


Fig 4. Implementation of Gating using a hybrid DBN classifier

The peaks obtained from the Pan Tompkins and by applying filters and machine learning are compared and if both of them state the same value of time at which peak occurs then the gating stops the alarm from being functioned. This uses an AND gate for the purpose of gating. The truth table for this gate is as constructed below.

TRUTH TABLE FOR AND GATE

Input		*Output*
A	***B***	
0	0	1
0	1	1
1	0	1
1	1	0

Performance Study
The following measures are being widely used in performanceevaluation [7].
Sensitivity (SEN), also well-known as hit rate or true positive rate(TPR), processes the proportion of correctly identified positivecases (sequences correctly classified as AFib) in regards tothe actual number of positive cases (all sequences identified as AFib) and is calculated by (6).

$$\text{SEN}=\frac{TP}{TP+FN} \quad (6)$$

Specificity, calculated by (7), measures the proportion ofnegative cases that are correctly classified (sequences correctlyclassified as not AFib) versus the total number of negativecases (all sequences identified as not AFib) in the observedsets.

$$\text{SPC}=\frac{TN}{TN+FP}(7)$$

Positive Predictivity Value (PPV), also known as precision,is the proportion of positive results that are true positives(sequences correctly classified as AFib) in regards to thetotal number of positive results (both true and false positives)acquired with the detection method, calculated by (8).

$$\text{PPV}=\frac{TP}{TP+FP}(8)$$

Accuracy (ACC) is calculated by (9) as the proportionof correct results including both the true and false positives(correctly classified sequences) in regards to the total number of sequences analyzed by the detection method.

$$\text{ACC}=\frac{TP+TN}{TP+TN+FP+FN}(9)$$

A statistical measure of a test's accuracy that combines SENand PPV is known as F1 score. The F1 score, also called Fscore or F Measure, is calculated by (10) and it represents theharmonic mean of the precision and sensitivity.

$$\text{F1} = \frac{2*TP}{2*TP+FP+FN}(10)$$

Of all these we have computed only the accuracy and is of major importance for the study of a hybrid DBN method gating.

RESULTS AND DISCUSSION

The results usually focus on enhancement of peaks using advanced filters for the minimum and maximum peaks being marked in detail.

Fig 5 depicts the maximum peaks obtained Fig 6 presents the minimum peaks obtained Fig 7 shows the way the maximum peaks are enhanced using the filtering process and echo peak detection algorithm in which the echo peaks are a prime reason for false alarms in our study.

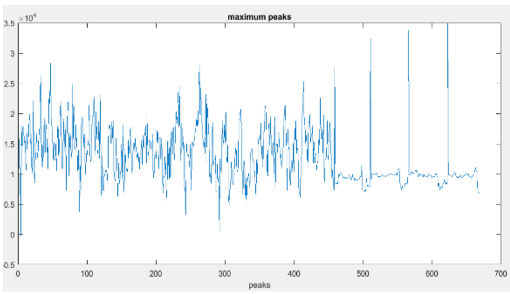
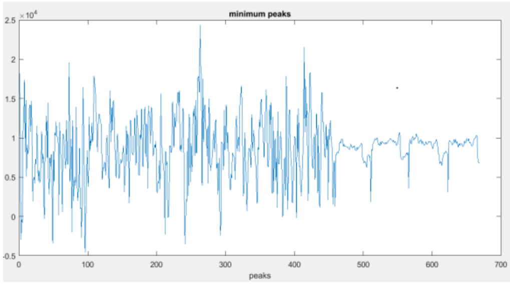


Fig 5. Maximum Peak detection



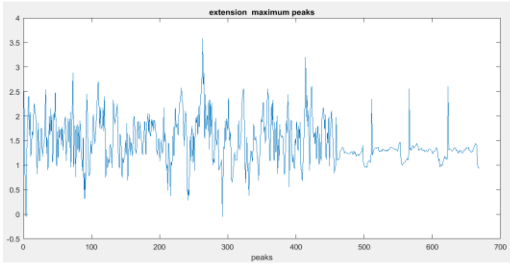


Fig 6. Minimum Peak De

Fig 7. Enhanced Echo Peaks in the Segments of ECG

Many other classifiers also underwent many tests, for instance the kNN classifier and Neural Network which have been proved to be impossible for classifying NSR and AFib on the scatter plots. Also the specificity was observed to be 0 %. Also another classifier namely Decision Tree classifier which has a specificity of 20 % also failed the test as its accuracy stated just over 83 % [8], [12].

Also for detecting AFib, two non linear statistical techniques have been implemented which involves Poincare plot and Sample Entropy [5] .But three statistical measured algorithms [4] have been used for good detection of peaks and their location. However this proved to be the best combination for the DBN classifier to detect and eliminate false alarms.

Deep learning algorithms have been grasping a bunch of consideration due to their ability to accurately perform problematical and computational tasks, especially when it comes to big and complex data [9]. A DBN is a deep learning algorithm thatis proficient to carry out the feature extraction segment of the learningprocess all by itself unlike other features which make the process more complex and hard to execute in case of large recordings. This is an eye-catching quality as featureextraction can be a time consuming and costly process. In theperspective of biomedical signals, DBNs have been applied toextort characteristics from six ECG signal types for arrhythmia classification [14].

Mere knowledge from deep learning in cardiology is enough for a professional to operate without any discomfort [13]. In [6], the viability of making use of a DBN is verified to differentiate clean and noisy ECG segment at five levelof signal-to-noise ratio. DBN, to a large extent, is a multilayer neural network which is made of quite a few layers of RBMs [36], [4].

In a standard neural network that has quite a lot of hidden layers, there is constantly the danger of being ensnared in the local minima. In addition, in such neural networks, training is a long process. DBN is asubstitute to overcomeall such non efficiencies and disadvantages in neural network operation. Training DBNs can be obtained by means of a layer-wise greedy algorithm. RBMs are trained layer by layer, and then a fine-tuning process is performed to adjust all the parameters of DBN [26].

Peaks identification is of major importance here which must be done in a very accurate manner. Echo peaks are the one which are generated due to noise in the surrounding area and for this hybrid DBN and echo peak detection algorithm is made use of and an accuracy of 85 % is achieved.

CONCLUSION

A new method having hybrid DBN architecture for a precise and accurate AFib detection is successfully developed in this paper. It is also aided to reduce false alarms during the patient surveillance. Machine learning is gaining a lot of attention now-a-days which is a well known fact. Many other algorithms have been developed having better accuracy than this but they have their corresponding drawbacks as discussed earlier and so this method is easy and can handle complex data in a very prompt manner and also the user can easily employ this with knowledge in the field of machine learning.

Furthermore, one can work in bringing out the most out of machine learning methods to enhance the accuracy still to a greater level but in an advantageous manner especially in the medical field as it is notorious that this field is a bit sensitive and needs a lot of attention to keep livelihood of people and minds of professionals at peace.

REFERENCES

[1] CDC - Centers for Disease Control and Prevention. (2017) Atrial fibrillation fact sheets. [Online]. Available: https://www.cdc.gov/dhdsp/data statistics/fact sheets/fs atrial fibrillation.htm.

[2] S. T. Mathew, J. Patel, and S. Joseph, “Atrial fibrillation: mechanistic insights and treatment options,” *European journal of internal medicine*, vol. 20, no. 7, pp. 672–681, 2009.

[3] *PhysioNet*. Accessed: Oct. 2016. [Online]. Available:https://www.physionet.org/physiotools/wag/nst-1.htm

[4] B. Taji, A. D. C. Chan and S. Shirmohammadi, "False Alarm Reduction in Atrial Fibrillation Detection Using Deep Belief Networks," in *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 5, pp. 1124-1131, May 2018.

[5] R. Mabrouki, B. Khaddoumi and M. Sayadi, "Atrial Fibrillation detection on electrocardiogram," *2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, Monastir, 2016, pp. 268-272.

[6] B. Taji, A. D. C. Chan, and S. Shirmohammadi, “Classifying measuredelectrocardiogram signal quality using deep belief networks,” in *Proc.IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, May 2017, pp. 1–6.

[7] M. Gusev and M. Boshkovska, "Performance Evaluation of Atrial Fibrillation Detection," *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, 2019, pp. 342-347.

[8] D. Altmann, K. Andler, K. Bruland, N. Nakicenovic and A. Nordmann, Converging Technologies - Shaping the Future of European Societies, 2004, [online] Available: http://bookshop.europa.eu/en/converging-technologies-pbKINA21357/.

[9] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, “Traffic flow predictionwith big data: A deep learning approach,” *IEEE Trans. Intell. Transp.Syst.*, vol. 16, no. 2, pp. 865–873, Apr. 2015

[10] .E. Farago and A. D. C. Chan, "Simulating Motion Artifact Using an Autoregressive Model for Research in Biomedical Signal Quality Analysis," *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, 2020, pp. 940-943.

[11] P. Bizopoulos and D. Koutsouris, "Deep Learning in Cardiology," in *IEEE Reviews in Biomedical Engineering*, vol. 12, pp. 168-193, 2019. S. Mahmood and I. A. M. Abdelrahman, "A Comparison between Different Classifiers for Diagnoses of Atrial Fibrillation," *2019 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*, Khartoum, Sudan, 2019, pp. 1-6.Qayyum, F. Meriaudeau and G. C. Y. Chan, "Classification of Atrial Fibrillation with Pre-Trained Convolutional Neural Network Models," *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, Sarawak, Malaysia, 2018, pp. 594-599.

[12] M. Huanhuan and Z. Yue, “Classification of electrocardiogram signalswith deep belief networks,” in *Proc. IEEE 17th Int. Conf. Comput. Sci.Eng. (CSE)*, Dec. 2014, pp. 7–12.

[13] K. Tateno and L. Glass, “Automatic detection of atrial fibrillation usingthe coefficient of variation and density histograms of RR and _RRintervals,” *Med. Biol. Eng. Comput.*, vol. 36, no. 6, pp. 664–671,Nov. 2001.

[14] D. Duverney*et al.*, “High accuracy of automatic detection of atrialfibrillation using wavelet transform of heart rate intervals,” *Pacing Clin.Electrophysiol.*, vol. 25, no. 4, pp. 457–462, Apr. 2002.

[15] S. Dash, K. H. Chon, S. Lu, and E. A. Raeder, “Automatic real timedetection of atrial fibrillation,” *Ann. Biomed. Eng.*, vol. 37, no. 9,pp. 1701–1709, Sep. 2009.

[16] J. Pan and W. J. Tompkins, “A real-time QRS detection algorithm,” *IEEETrans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.

[17] H. Sedghamiz. *Complete Pan Tompkins Implementation ECG QRS Detector*. [Online]. Available: http://www.mathworks.co.kr/matlabcentral/fileexchange/45840-complete-pan-tompkinsimplementation-ecg-qrsdetector.

[18] Benjamin, E. J., P. A. Wolf, R. B. D’Agostino, H.Silbershatz, W. B. Kannel, and D. Levy. Impact of atrialfibrillation on the risk of death: the Framingham HeartStudy. Circulation 98:946–952, 1998.

[19] Heeringa, J., D. A. van der Kuip, A. Hofman, J. A. Kors,G. van Herpen, B. H. Stricker, T. Stijnen, G. Y. Lip, andJ. C. Witteman. Prevalence, incidence and lifetime risk ofatrial fibrillation: the Rotterdam study. Eur. Heart J. 27:949–953, 2006.

[20] Wolf, P. A., R. D. Abbott, and W. B. Kannel. Atrialfibrillation as an independent risk factor for stroke: theFramingham Study. Stroke 22:983–988, 1991.

[21] H. Bozdogan, “Model selection and Akaike’s Information Criterion(AIC): The general theory and its analytical extensions,” *Psychometrika*,vol. 52, no. 3, pp. 345–370, Sep. 1987.

[22] M. A. Carreira-Perpinan and G. E. Hinton, “On contrastive divergencelearning,” in *Proc. Artif. Intell. Statist.*, vol. 10. 2005, pp. 33–40.

[23] R. Hecht-Nielsen, “Theory of the backpropagation neural network,” in*Proc. Int. Joint Conf.*, 1989, pp. 593–605.

[24] G. E. Hinton, S. Osindero, and Y.-W. Teh, “A fast learning algorithmfor deep belief nets,” *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554,2006.