ATILIM UNIVERSITY

GRADUATE SCHOOL
OF
NATURAL AND APPLIED SCIENCES
INSTRUCTIONS FOR PREPARATION

OF THESES

DEEP LEARNING APPLICATIONS IN SEIZURE DETECTION

A MASTER'S THESIS

In

Electrical and Electronics Engineering

Atilim University

By

YEZI ALI KADHIM

JUNE 2018

DEEP LEARNING APPLICATIONS IN SEIZURE DETECTION

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF NATURAL AND

APPLIED SCIENCES

OF

ATILIM UNIVERSITY

BY

YEZI ALI KADHIM

IN PARTIAL FULFILLMENT
OF
THE REQUIREMENTS FOR THE DEGREE
OF

MASTER

IN
THE DEPARTMENT OF

ELECTRICAL AND ELECTRONICS ENGINEERING

JUNE 2018

Approval of the Graduate School of Natural and Applied Sciences, Atilim University.

_____

(Prof. Dr. Ali Kara)

Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

_____

(Assoc. Prof. Dr. Efe Eseller)

Head of Department

This is to certify that we have read the thesis "Deep learning applications in Seizure detection" submitted by "Yezi ali kadhim" and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

(Associate Prof. Dr. Lalit Garg)	(Prof. Dr. Alok Mishra)
Co-Supervisor	Supervisor

Examining Committee Members

Prof. Dr. Erdoğan Doğdu _____

Cankaya University – Computer Engneering

Prof. Dr. Alok Mishra _____

Atilim University – Software Engneering

Associate Prof. Dr. Reşat Doruk _____

Atilim University – Electrical and Electronics Engneering

Date: 18.06.2018

I declare and guarantee that all data, knowledge and information in this document has been obtained, processed and presented in accordance with academic rules and ethical conduct. Based on these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Yezi Ali Kadhim

Signature:

ABSTRACT

Deep learning applications in Seizure detection

Kadhim, Yezi Ali

M.S., electric and electronic Engineering Department

Supervisor: Prof. Dr. Alok Mishra

Co-Supervisor: Assoc. Prof. Dr. Lalit Garg

June 2018, 52 pages

In this thesis, a new method is proposed based on deep auto encoder and power spectral density. First, the input data is analyzed using power spectral density for feature extraction by measuring the power spectral density of the signal for each row of data. The produced output becomes input to the first Auto encoder to reduce the dimension and extracted high level features. The output of first auto-encoder become input to the second auto-encoder also to reduce number of features and extracted high level features. In addition, these features are classified into two groups: normal and abnormal by using SoftMax classifier. Finally, the two auto-encoders and SoftMax stacked and trained by using backpropagation algorithm to improve the classification accuracy. The proposed method gives satisfactory results when compared with the common methods presented in this filed .Here, the number of Auto encoders depend on the behavior of the data as well as the dimension. The proposed method is tested with commonly used datasets in the epilepsy serius detection, and the results obtained are compared with other and most prominent works in this field in order to determine the strengths and weaknesses.

Keywords: Deep Auto encoder, power spectral density, epilepsy serius detection

ÖZ

Nöbet tespitinde derin öğrenme uygulamaları

Kadhim, Yezi Ali

Yüksek Lisans, Elektrik-Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Prof. Dr. Alok Mishra

Ortak Tez Yöneticisi: Assoc. Prof. Dr. Lalit Garg

Haziran 2018, 52 sayfa

Bu tezde, Derin otomatik kodlayıcı ve güç spektral yoğunluğuna dayalı yeni bir yöntem önerilmiştir. ilk, giriş verileri, her bir veri satırı için sinyalin güç spektral yoğunluğunu ölçerek özellik çıkarımı için güç spektral yoğunluğu kullanılarak analiz edilir. Üretilen çıktı, boyutu azaltmak ve yüksek seviyeli özellikler elde etmek için ilk Otomatik kodlayıcıya girdi olur. İlk otomatik kodlayıcının çıkışı, özelliklerin sayısını azaltmak ve yüksek seviyeli özellikler elde etmek için ikinci otomatik kodlayıcıya girdi olur. Ayrıca, bu özellikler iki gruba ayrılır: SoftMax sınıflandırıcı kullanılarak normal ve anormal. Son olarak, iki otomatik kodlayıcı ve SoftMax, sınıflandırma doğruluğunu geliştirmek için geri yayılım algoritması kullanılarak yığılmış ve eğitilmiştir. Önerilen yöntem, bu dosyada sunulan ortak yöntemlerle karşılaştırıldığında tatmin edici sonuçlar verir. Burada, Otomatik kodlayıcıların sayısı verilerin yanı sıra boyutun davranışına da bağlıdır. Önerilen yöntem, epilepsi seri tespitinde yaygın olarak kullanılan veri setleri ile test edilmiş ve elde edilen sonuçlar, güçlü ve zayıf yönleri belirlemek amacıyla bu alandaki diğer ve en önemli çalışmalarla karşılaştırılmıştır.

Anahtar kelimeler: Derin Otomatik Kodlayıcı, güç spektral yoğunluğu, epilepsi seri saptama

To My Parents

My Father “ ALI KADHIM ISAMIL”

My Mother “ATHRAA ABDULGHAFOOR IBRAHIM”

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere appreciation to my supervisor Prof. Dr. Alok Mishra for his guidance, support. Patience and insight throughout my thesis study. I offer sincere thanks for his extensive knowledge, incentive and patience during this period. I am really proud of being one of his student.

I would also express my appreciation to examination committee members, for their valuable time, comments and suggestions.

Also I would like to thank my co-supervisor Assoc. Prof. Dr. Lalit Garg for all support.

Last but not least, I would like to express my appreciation to my father, Ali Kadhim Ismail , and my mother, Athraa AbdulGhafoor Ibrahim for their endless love, support and patience through the study and all my life .

TABLE OF CONTENTS

ABSTRACT iii

OZ iv

DEDICATION..... v

ACKNOWLEDGMENTS vi

TABLE OF CONTENTS vii

LIST OF TABLES ix

LIST OF FIGURES x

LIST OF ABBREVIATIONS..... xi

CHAPTER

1. INTRODUCTION 1

1.1 Power Spectral Density 2

1.2 Auto-encoder 3

1.3 Objective of Study 5

1.4 Motivation..... 5

1.6 Research Questions..... 5

2. BACGROUND INFORMATION AND LITERATURE SURVEY 5

2.1 Support Vector Machine (SVM) 6

2.2 Naive Bayes algorithm (NB)..... 8

2.3 M ulti-Layer Perceptron (MLP) 8

2.4 Radial Basis Function Network (RBF) 9

2.5 Decision Tree (DT) 11

2.6 Related Works..... 11

3. MATERIAL AND METHODS..... 24

3.1 Deep Learning 24

3.1.1 Recurrent Neural Network (RNN) 25

3.1.2 Convolutional neural network (CNN) 27

3.1.3 Deep Neural Network (DNN) 29

3.2 Power Spectral Density (PSD) 29

3.3 Proposed Method 30

4. RESULTS 33

5. DISCUSSION OF COMMON ATTRIBUTES AND CONCLUSION 36

5.1 Discussion..... 36

5.2 Conclusion..... 36

5.3 Limitation of Study 37

5.4 Future Work 37

REFERENCES 38

LIST OF TABLES

TABLE

1. Comparison of Results with Previous Works.....35

LIST OF FIGURES

FIGURES

1. Power Spectral Density..... 2
2. Green and red dot classification..... 6
3. Nonlinear Model..... 7
4. Classification nonlinear model using SVM..... 7
5. Multilayer layer Perceptron..... 9
6. Radial Basis Function Network..... 10
7. Simplest form of fully recurrent neural network..... 26
8. Simple Recurrent Networks..... 27
9. Convolutional Neural Network..... 28
10. Deep neural network based on PSD for Epileptic seizure..... 32
11. PSD of Train Data for Epileptic seizure (Normal and Abnormal)..... 33
12. PSD of the Test Data for Epileptic seizure (Normal and Abnormal)..... 34
13. Results of deep Auto encoder based on PSD..... 35

LIST OF ABBREVIATIONS

SVM	-	Support Vector Machine
NB	-	Naive Bayes algorithm
RBF	-	Radial Basis Function Network
DT	-	Decision Tree
DNN	-	Deep Neural Network

1. INTRODUCTION

Epilepsy is a nervous malfunction affecting around 1-2% of all humans worldwide [1]. It is distinguished by surprised repeated and evanescent troubles of insight or behavior produce from extreme harmonization of cortical neuronal networks; it is a nervous disorder in which a person expertise chronic irregular explosion of electrical releases in the brain. The lineament of epilepsy is repeated crisis labelled "epileptic crisis ". Observation brain performance over the electroencephalogram (EEG) has become a serious tool in the detection of epilepsy [65]. The EEG records of patient's hardship from epilepsy display two types of abnormal activity: inter-ictal, abnormal signals verified between epileptic crisis; and ictal. The EEG sign of an inter-ictal action is episodic transitory waveforms, as either secluded spikes, spike trains, acute waves or spike-wave complexes. EEG sign of an epileptic crisis is unruffled of a continuous ejection of polymorphic waveforms of mutable largeness and frequency, spike and shrill wave campuses, rhythmic hyper synchrony, or electro brainy idleness observed over a period slower than the average period of these irregularities through inter-ictal stages [66, 2].

Commonly, the recognition of epilepsy can be accomplished by visual survey of EEG records for inter-ictal and ictal actions by a veteran neurophysiologist. However, visual survey of the huge extent of EEG information has businesslike weaknesses. Visual search is very time overwhelming and inactive, essentially in the situation of long-term records. In other hand, dispute among the neurophysiologists on the similar records is potential due to the individual nature of the examination and due to the set of inter-ictal spikes morphology. Furthermore, the EEG designs that describe an epileptic crisis are comparable to waves that are fragment of the contextual noise and to objects such as eye winks and another eye motion, muscle action, electrocardiogram, conductor "pop" and electrical meddling. For these purpose, techniques for the automatic recognition of inter ictal spikes and epileptic crisis can avail as worthy medical tools for the scanning of EEG records in an extra topical and computationally active method.

In this thesis, we devise a deep neural network based on power spectral density (PSD) with the main features extracted by means of using PSD applied to EEG signals and, then, classifying those features using deep neural network.

The proposed method presented satisfactory results when compared with most famous studies in this field.

1.1 **Power Spectral Density (PSD)**

In this section, we will briefly explain how the strength of a signal is divided in the frequency domain proportional to the strengths of any signals in the environment. This is key in terms of the design features of all linear time-invariant filters intended to quell or select a signal [4]. This feature is ideal when signals are deterministic, and proves to be the right alternative when it comes to random signals. For example, if a metrical waveform is an audio signal within collective disorder signals, it becomes necessary to create a low-pass linear time-invariant filter so as to extract the audio signals and separate the disorder signals [5][6]. The PSD function presents the force of the energy in the signal as a function of frequency. The unit of this function is energy (expressed in terms of difference) for each frequency (expressed in terms of width), and can gain energy in a certain frequency domain by means of merging the PSD in that frequency domain. Figure 1 is an example of PSD signal.

Fig 1: Power Spectral Density

1.2 Auto-encoder

The greedy layer sage method for pre-training a deep auto-encoder neural network works by training each auto-encoder, which represents a layer in turn. In our study, auto-encoders are initially trained in the form of unsupervised learning [7].

A stacked auto encoder (SAE) is a neural network generated from multiple layers of sparse auto-encoders once the outputs of each layer (auto-encoder) turn into inputs for the next layer. In the mathematical sense, in order to represent the SAE, the number of auto-encoders is represented as n layers [8][9]. Let $W^{(k,1)}$, $W^{(k,2)}$, $b^{(k,1)}$, $b^{(k,2)}$ indicate the parameters $W^{(1)}$, $W^{(2)}$,$b^{(1)}$,$b^{(2)}$ for the kth auto-encoder. The encoding level for the SAE is given by implementing the encoding stage of each layer in forward order:

$$a^{(t)} = f\,(z^{(l)}) \tag{1}$$

$$z^{(t+1)} = W^{(l,1)} \quad a^{(l)} \quad + b^{(l+1)} \tag{2}$$

Fig 2: Auto-encoder

The decoding level is specific by executing the decoding with the stacked auto-encoder in inverse order:

$$a(n+l) = f\,(z(n+l)) \qquad (3)$$

$$z(n+l+1) = W(n-l,2)\quad a(n+l)\quad + b(n-l,2) \qquad (4)$$

The input of the benefit is included within a(n), which is the activation of the deepest layer of the hidden parts. This array provides a good example of of the information in terms of high-level features. By passing a(n) to a SOFTMAX function, the features

produced by the stacked auto encoder can be used for classification problems to deal with problems related to classification [10].

1.3 Objective of study

The present thesis attempts to develop a new framework so as to combine the features of deep learning and power spectral density for the automatic detection of epilepsy serious. The proposed framework will be validated with the help of common datasets in this field, and the results will be compared with the best methods presented in the literature.

1.4 Motivation

Up to the present, the applications of deep learning in other fields of research has yielded satisfactory results; these include image recognition, disease detection and data classification. Deep auto-encoder, for the purpose of detection, presents only 60% accuracy which is too low. However, combined with PSD, the encoder will have increased accuracy, which is the main novelty of the present study.

1.5 Research Questions

In this section, a number of questions is presented as research questions to be addressed by the study:

- How can we determine the best detection method for epilepsy serious detection?
- How can the accuracy of the best determined features and classifiers be increased?
- How can multiple detection methods be tested for epilepsy serious detection, and how can the results be best evaluated?

2.REALATED WORK

2.1 Support Vector Machines (SVM)

Support Vector Machines rely on the idea of decision seaplanes that describe choice limits. A decision smooth is one that splits among a group of items with various class memberships. A diagram instance is presented in Figure 2.2. In this instance, the items belong to classes GREEN or RED (change all to green and red). The splitting line describes a limit on the right of which all items are GREEN and to the left of which all items are RED. Any different item (white circle) dropping to the right is branded, i.e., categorized, as GREEN (or categorized as RED should it drop

Figure 2.2: Green and red dot classification [11]

to the left of the splitting line. The beyond is a standard sample of a linear classifier, i.e. one that splits a group of items into their own sets (GREEN and RED in this situation) with a line. However, the most practical classification errands are not that humble, and regularly more composite organizations are required in order to create ideal separations and appropriately categorize new items (or testing) within the obtainable samples (or training). This state is shown in Figure 2.3. Compared to the preceding diagram in figure 2.3, it is likely that a full separation of the GREEN and RED items would need a curve. Classification errands depend on sketch splitting lines to differentiate among items of several classes. In addition, their associations are identified as hyperplane classifiers. In this respect, Support Vector Machines are mainly suited to handle such errands.

Figure 2.3: nonlinear model [12]

Figure 2.4 displays the simple impression behindhand Support Vector Machines. Here, we see the innovative items drawn, i.e., reordered, using a group of mathematical functions and identified as kernels. The procedure of reorganizing the substances is identified as charting (that is, conversion). Note that, in this specific scenario, the recorded items on the right side of the diagram are linearly separate instead of building the compound curve on the left side of the diagram. Hence, a most suitable curve is formed so as to split the GREEN and the RED items.

Figure 2.4: Classification nonlinear model using SVM [13]

2.2 Naive Bayes algorithm

It is a classification method based on Bayes' Theorem with the idea of individuality amongst the predictors. In basic terms, a Naive Bayes classifier accepts that the appearance of a specific feature in certain classes is not related to the appearance of any other feature. To set an example, a fruit may be regarded as an apple if it is red, round, and about 3 creeps in length. Even if these features rely upon each other or upon the presence of other certain features, all together they can lead us to the notion that this fruit is an apple, hence the term 'Naive'. Naive Bayes model is informal to form and primarily applicable to very large data groups. Along with minimalism, Naive Bayes is considered to even surpass extremely high-level and recognized classification approaches.

2.3 Multi-Layer Perceptron (MLP)

The area of artificial neural networks (ANN) has been quite often referred to as merely neural networks, in which multi-layer perceptron is perhaps the most commonly recognized feature. A perceptron is an individual neuron model as a forerunner to greater neural networks. ANN examines how basic models of biological brains can be used to accomplish various computations and tasks, such as prognostic modeling jobs in machine learning. Here, the purpose is not to make a replica of the brain and, rather, to improve healthy algorithms and data structures to be later used for modeling potential problems.

The main function of neural networks lies in the capacity to learn the features in the training data and how to apply them in the most efficient way to the output variable that is to be predicted. In this intelligence neural networks learn a representing. In theoretical terms, any plotting function can be learnt and established as a general estimate algorithm.

The prognostic ability of neural networks comes from the ordered or multi-layered organization of the networks themselves. As shown in Figure 2.5, the structure of MLP input layer, more than one hidden layer and output layer represent the predication value. The output neurons represent the number of predication and the number of neurons in the input and the hidden layer represent the number of features.

Figure 2.5: Multilayer layer Perceptron [13]

2.4 Radial Basis Function Network (RBF)

Radial functions can work in any type of systems, whether linear or non-linear, and any type of network, single-layer or multi-layer. Though, based on Broom and Lowe's 1998 seminal paper, RBF networks are commonly related with radial functions in a single-layer network, such as the one shown in Figure 2.5.

In an RBF non-linear network, the bias functions can change or move size and there is more than one hidden layer. Here, our focus will be on single-layer networks with functions which are fixed in position and size.

Figure 2.5: Radial Basis Function Network [14]

2.5 Decision tree

Decision trees are a kind of supervised machine learning mechanisms, where the information splits on a non-stop basis and according to a fixed parameter. The tree is characterized by two components, namely decision nodules and greeneries. The greeneries are the decisions or the last consequences. Also, the decision nodules are where the information splits.

There are two major kinds of decision trees. These are; classification trees, or Yes/No type; for instance, where the consequence is a parameter like ‘fit’ or ‘unfit’. Here the result parameter is categorical and not continuous. The other form is the regression tree, or continuous data tree, where the decision or the output results are continuous, e.g. a number like 1234.

2.6 Related Works

In this section, number of previous studies will be presented to compare results methods that used to detect epilepsy Seizures. In [15] Ralph G. Andrzeja et al. compared dynamic feature of brain signal action from various recording parts and from various physiological and pathological brain situation. Nonlinear prognosis mistake used and a predicate of an active connexon dimensions in group with the mechanism of courteous capacity adjust replacement input, they analyzed groups of EEG! time chain: roof EEG enrollment from healthy recruit with closed eyes shut and opened eyes, and intracranial EEG enrollment from epilepsy patient through the crisis permitted period from leaving the crisis sire region as well as cranial EEG enrollment of epilepsy crisis.

As a pre-analysis stage a modulation principle of prosaic stationarity was used. roof EEG enrollment with eyes exposed were appropriate with the replacement’ void supposition of a Gaussian linear stochastic(GLS) procedure. In [16] V. Srinivasan et al. explain an automated technique for diagnosis epileptic crisis using a specific kind of recurrent neural network(RNN) recognized as Elman network.

The trial is hold out by employ time domain in addition to frequency domain EEG signals feature. empirical effects present that Elman network crop epileptic recognition performance amounts as satisfactory as 99.6% with a particular data attribute which is well than the outcomes gained by using further kinds of neural networks with two and or input attribute. In [17] Abdelhamid Subasi et al. introduce two principally diverse methods to design classifiers the classical statistical techniques depended on logistic regression (LR) and the emergent computational influential methods depended on artificial neural network. LR in addition MLPNN based classification modules were sophisticated and contrast in connection to their accuracy in EEG signals classification.

In these approaches they recycled LBDWT factors of EEG signal as data to recognition method with two separate productions: epileptic crisis (abnormal) or non-epileptic crisis (normal). By recognizing feature in the signal, they want to offer an involuntary method that will provision a doctor in the identifying procedure. By relating LBDWT in linking with MLPNN, they gained new and dependable classifier design. The contrasts between the industrialized classifiers were principally based on examination of the receiver operational specific (ROC) curves as well as amount of scalar accuracy. In [18] Abdelhamid Subasi et al. presented Combination of experts (ME) is integrated NN design for supervised learning.

A dual - loop Expectancy - Maximization (EM) technique has been announced to the ME net organization for recognition of epilepsy crisis. The recognition of epileptiform releases in the EEG is a significant module in the analysis of epilepsy. EEG signals were disintegrated into the frequency sub-bands using separate discrete wavelet transform (DWT).

Then, these sub-band frequencies which used as a data to a ME net with two separate productions: normal and abnormal. In order to increase performance, the productions of professional systems were mutual rendering to a group of resident weights termed ‘‘gating function’’.

The invariant conversions of the ME possibility density functions contain the variations of the practiced tags and the conversions of the factors in the gating functions. The accuracy of the presented system was assessed in relations of recognition precisions and the effects verified that the suggested ME net organization

has some possible in recognition epilepsy crisis. The ME net organization accomplished performance degrees which were greater than that of the impartial NN model. In [19].

Kannathalet al proposed the electroencephalogram (EEG) is an illustrative signal covering data around the situation of the brain. The form of the wave may cover appropriate data about the general of the brain. Though, the human viewer cannot openly display these understated parts. Furthermore, since bio signals are extremely individual, the indications may seem at accidental in the time scale. Hence, the EEG signal factors, mined and evaluated using machines, are very valuable in diagnostics. The purpose of this research is to compare the various entropy predictor when smeared to EEG data from abnormal and normal issues.

The results achieved show that entropy estimators can separate normal and epileptic EEG cases with more than 95% sureness. The recognition capability of the entropy methods is verified using ANFIS classifier. The results are talented and a recognition accuracy of about 90% is succeeded.

In [20] The footage of attack is of major concern in the estimation of epilepsy cases. crisis is the spectacle of rhythmicity ejection since any a native part or the entire brain and the single actions typically continues from seconds to minutes. Then crisis, in overall, happen uncommonly and randomly, automatic recognition of crisis through long-term EEG footages is extremely suggested.

As EEG signals are non-stationary, the conservative approaches of frequency study are not positive for recognition determinations.

The author shows a technique of study of EEG signal, which is depended on time frequency investigation. Originally, determined sections of the EEG signals are investigated using time frequency approaches and numerous attributes are mined for each section, representative the energy spreading in the time frequency flat. Formerly, those attributes are used as a data in ANN, which offers the last recognition of the EEG sections regarding the being of crisis or not. The researcher used a widely accessible dataset in command to calculate proposed method results and the assessment results are very hopeful representative general performance from 97.72% to 100%. In [21]

Kemal Plat et al. presented study to classify epilepsy seizure in EEG signals using a mixture system depended on decision tree classifier and fast Fourier transform (FFT).

The current research presents a hybrid structure with two steps: attributes selection by FFT and classification by DT classifier. The recognition of epileptic form releases in the EEG is an imperative chunk in the recognition of epileptic. Overall dataset was gotten from EEG signal of normal topics and topics sorrow from epilepsy crisis. For healthy topics is related EEG (scalp) with uncluttered eyes and for abnormal patients resemble to a seizure logged in hippo-campus (epileptic effort) with depth rods.

The development of presented method was displayed using 5-fold cross-validation, recognition performance, and sensitivity and specificity standards. they have gotten 98.68% and 98.72% cataloging precisions using 5- and 10-fold cross-validation. The specified consequences display that the suggested technique could point out the ability of scheme of a new intelligent support analysis structure. In [22] U. Rajendra Acharya et al. proposed research the appearances of epileptic events. Owed to the dynamic landscape and non-linear of EEG signals, visual review and reading of these signals are deadly, time-consuming, error-horizontal, and exposed to inter-witness inconsistencies. Consequently, numerous Computer Aided Diagnostic (CAD) established educations have accepted non-linear methods to investigate the abnormal and normal in EEGs.

The researcher proposed a new automatic method depended on feature excavating for epilepsy action recognition. In order to associate this research with the consequences of relation researches in the survey, they used the commonly used standard database from Bonn University for valuation of presented method. Hundreds examples each in normal, interictal, and ictal classes were used. they disintegrated these sections into wavelet coefficients using Wavelet Packet Decomposition (WPD) and mined eigenvalues commencing the subsequent wavelet coefficients using PCA.

Important eigenvalues, determined using the ANOVA test, were recycled to train and validated various supervised learning using the 10-fold applied cross validation method. researcher gained 99% detection performance using the Gaussian Mixture Model (GMM) classifier. The presented method is accomplished of detecting EEG

sections with clinically acceptable accuracy using fewer number of features that can be extracted with fewer computational cost.

The method can be transcribed as a software claim that can be simply organized at a little rate and used with virtually no knowledgeable training. they expect that this application can, in the upcoming, develop into an effective assistant instrument that cannot only recognize epilepsy events in EEG signals but also repeatedly display the onset of attacks and thus assistance the physicians in as long as improved and timely upkeep for the patient's sorrow from epileptic.

In [23] U. Rajendra Acharya et al. developed a Computer Aided recognition method with least preprocessing stages that can recognize all the three sessions of EEG sections, namely abnormal, normal, and ictal using a minor numeral of very discerning nonlinear attributes in humble classifiers. To assess the method, sections of normal, abnormal, and ictal sections were used. Nonlinear attributes depended on the Upper Order Bands, two entropies, namely the Estimate Entropy and the Example Entropy, and Fractal Measurement and Hurst Advocate were mined from the sections. Important attributes were determined using the ANOVA test.

After assessing the enactment of six classifiers (Fuzzy Sagano Classifier, K-Nearest Neighbor, Gaussian Mixture Model, Support Vector Machine, Decision Tree, Radial Basis Probabilistic Neural Network,) using a mixture of the determined features, they found that using a group of all the determined six attributes in the Fuzzy classifier resulted in 99.7% recognition precision. they have verified that proposed method is talented of reaching high precision using a minor number of features that correctly detention the understated various in the three various kinds of EEG (normal, abnormal, and ictal) sections.

The method can be simply inscribed as a software presentation and used by medical specialists without any general training and cost. Such application can grow into an involuntary crisis nursing application in the near future and can assistance the physicians in supplying improved and timely care for the patient's sorrow from epileptic.

In [24] Musa Piker et al. proposed a novel technique for the identification of epilepsy from EEG signals depend on multifaceted classifiers. To transmit out this revision,

first the attributes of EEG signal are mined by using a DWT at various stages of granularity to gain size drop. In following stages, five attribute (depended on statistical capacities maximum cost, minimum cost, mathematics nasty, normal nonconformity, average value) are acquired by using the attribute vectors and are obtainable as the input measurement to the complex-valued neural networks(CVNN). The assessment of the suggested technique is led using the k-fold cross-validation approach, reporting on recognition performance, specificity, and sensitivity.

The suggested technique is verified using a standard EEG database, and high correctness proportions were acquired. The specified consequences display that the suggested manner can be used to policy a correct recognition model for epilepsy diagnosis.

In [25] Nigam et al. describes a technique for schematize recognition of epilepsy crisis since EEG signals using a multilevel nonlinear preprocessing in mixture with an indicative (LAMSTAR) ANN. Preprocessing through multilevel non-linear. LAMSTAR data training, ANN training and classification accuracy (1.6% miss rate, 97.2% complete precision when considering equally false-alarms and 'misses') are conferred and are presented for comparing kindly with previous methods offered in current works.

In [26] Themis P. Exarches presented a method for the automatic recognition and recognition of fleeting actions in EEG records is offered. It is depended on connotation regulation withdrawal and recognize passing actions into four classes: epilepsy spikes, influence activity, eye irregular action, and shrill alpha action. The method includes four steps: 1) fleeting event recognition; 2) clustering of fleeting events and features mining; 3) features discretization and features subsection range; and 4) connotation rule excavating and recognition of fleeting events. The approach is estimated using 25 EEG footages, and the best-found precision was 87.38%. The offered method associations high accurateness with the aptitude to offer clarification for the results prepared, since it is depended on a set of connotation rules.

In [27] Faure C. presents an organizational pattern classification method for signals understood and labelled by a social practiced. For any shape to be known, the signal is characterized by a series of local data which is syntactically precise if some

'structural features' are content. Kinds of primitives and organizational features are well-defined. This demonstration technique takings a priori information of the specialist's explanations into account. the system is worried with EEG's analysis. Recognition of fleeting proceedings, points and spike-and-wave campuses assistances the electroencephalographs by refusing fragments of the EEG suggestion which surely do not cover these forms. The fleeting actions detection saves a 16% middling of the preliminary proposal for compulsive EEG's and 8.3% for non-pathological EEG's. Only 3.3% of pathological EEG's are not disallowed afterward points and spike-and-wave campuses recognition. The outcomes of this analyze are equated with human clarifications.

In [28] Hiram Firpi proposed a general-resolution, methodical algorithm, containing of a hereditary programming unit and a k-nearest neighbor classifier to repeatedly build artificial features right from the recreated state-space route of the EEG signals that disclose patterns prognostic of epileptic crisis. The algorithm was assessed in three altered patients, with reduction well-defined above a prospect of 5 minutes earlier clear electrographic onset. Trials are approved out using 20 starting point epochs (non-crisis) and 18 parietal epochs (pre- crisis). Outcomes display that just two crises were missed while a perfect detection on the starting point epochs was realized, soft a 0.0 false positive per hour. In [29] A.J. Gabor proposed Computerized crisis detection using a self-organizing neural network. Electrodense clin. Neurophysiology., 1996, 99: 257–266) was authenticated with 200 annals from 65 cases (4553.8 h of recording) covering 181 crises. Design and approaches: accuracy of the technique was obvious by its sensitivity ((crisis recognize/total crisis) ´ 100) and selectivity (false-positive errors/Hr-FPH). Contrasts with the Screen recognition technique and audio-transformation were done. Consequences: CNET recognize 92.8% of the crisis and had a nasty FPH of 1.35 ± 1.35. Screen recognized 74.4% of the crisis and had a mean FPH of 3.02 ± 2.78. Audio transformation detected all but 3 (98.3%) of the crisis. Selectivity for this recognition policy was not definite. Deductions: This study not only authenticates the CNET algorithm, but also the idea that crisis have frequency-amplitude topographies that are contained in signal space and can be selectively recognized as being different from other kinds of EEG forms. The auricle is a specific

frequency-amplitude sensor and when the signal is converted into audio frequency range, crisis can be recognized with improved sensitivity as equated to the other policies inspected. In [30] Gabor et al. presented Feed-forward, artificial neural networks with error-back-propagation were realistic to credit of epileptic system forms in the EEG.

The integral network feature of simplification and erraticism lenience were active in recognizing wave procedures that varied from the training forms but motionless preserved 'epileptiform' spatio-temporal appearances. The inevitability of credit was dignified as a incessant role with a variety of 0-1. Two stages of cert (0.825 and 0.900) were recycled to designate appreciation of points and shrill waves (SSW). An regular 94.2% (+/- 7.3) of the SSW were predictable; 20.9% (+/- 22.9) of all accepted SSW were false-positive classification. The time essential for pattern classification was fine inside the time necessary for digitizing the analogue information. This study offers indication that NN tools, in principle, a real pattern classification approach for empathy of epileptiform transients in the EEG.

The study is satisfactorily fast to be of possible rate as a policy for information reduction of extended footages stowed on unpackaged media. In [31] Amir B. Gave proposed an approach, EEG-depended, brain-case empathy technique which could usage the foundation for predicting a global epilepsy crisis. The technique trusts on the presence in the EEG of a premeasure case, with extractable single features, a priori approximate. We bare 25 pests to hyperbaric oxygen pending the arrival of a sweeping EEG crisis. EEG subdivisions from the preexposure, initial experience, and the epoch up to and containing the crisis were treated by the fast wavelet transform. features selected from the wavelet coefficients were recorded to the unsupervised optimum fuzzy clustering (UOFC) technique.

The UOFC is suitable for categorizing comparable sporadic sequential objects in the semi stationary EEG to a group of clusters which may denote brain-case. The unsupervised choice of the several of clusters overwhelms the a priori unidentified and mutable numeral of conditions. The typically unclear brain formal changeovers are logically preserved by transmission each sequential shape to one or more fuzzy clusters. The recognition prospered in recognizing some, comportment supported, EEG cases like as slumber, latent, attentive and lively restlessness, as well as the crisis.

In 16 examples a premeasure state, fixed between 0.7 and 4 min was distinct. Substantial separate erraticism in the number and features of the clusters may delay the awareness of an early general epileptic notice. Universality may not be critical if by means of a dynamic type of the UOFC which has been qualified the entity's standard terminology of EEG cases and can be predictable to recognize unnamed new cases.

In [32] A new wavelet-chaos-NN approach is offered for recognition of EEGs into normal, ictal, and abnormal EEGs. Wavelet study is used to crumble the EEG into delta, theta, alpha, beta, and gamma sub-sets. Three factors are working for EEG illustration: normal nonconformity, correlation dimension, and major Lyapunov advocate.

The recognition precisions of the subsequent methods are equated: 1) unsupervised-means clustering; 2) linear and quadratic discriminant examination; 3) radial basis function NN; 4) Levenberg–Marquardt backpropagation neural network (LMBPNN). To decrease the calculating period and production investigation, the research was achieved in two parts: band-specific part and combined-band part. In stage two, over 500 various combinations of combined-band feature seats involving of talented factors from stage one of the study were examined.

It is decided that all three important modules of the wavelet-chaos-neural network methods are essential for successful the EEG detection performance. Sensible mixtures of factors and classifiers are required to precisely distinguish between the three kinds of EEGs. It was exposed that a specific combined-band attributes interplanetary involving of nine factors and LMBPNN outcome in the maximum recognition performance, a maximum assessment of 96.7%.

In [33] Samanwoy Ghosh-Dastidar et al. proposed a new PCA-improved cosine RBFNN classifier is proposed. The two-levels classifier is combined with the combined-band wavelet-chaos method, industrialized previous by the authors, for precise and healthy recognition of EEGs into normal, ictal, and abnormal EEGs. A nine-factors combined-band attributes space exposed in earlier studies for active EEG illustration is used as information to the two-level classifier. In the first level, PCA is used for attributes improved. The reorganization of the data space lengthways the

principal components of the input develops the recognition performance of the RBFNN working in the second level significantly.

The recognition performance and healthiness of the classifier are authorized by general parametric and sensitivity investigation. The new wavelet-chaos-neural network approach produces high EEG detection performance (96.6%) and is fairly forceful to changes in training information with a small usual nonconformity of 1.4%. For epileptic detection, when one healthy and abnormal EEGs are measured, the detection performance of the suggested system is 99.3%. This measurement is exclusively extraordinary because even the greatest extremely trained neurologists do not seem to be capable to recognize interracial EEGs greater than 80% of the periods.

In [34] Glover JR Jr et al. defines a knowledge-depended model for the abolition of false positives in the computerized recognition of epileptogenic shrill transients in the EEG. The system sorts inclusive use of longitudinal and chronological situation data obtainable on 16 networks of EEG, EKG, EMG, and EOG.

A knowledge-depended execution is recycled because of the comfort with which it lets the related rules to be voiced and superior. The ensuing system is displayed to be accomplished of refusing a wide range of objects generally found in EEG records, objects that cause several false positive discoveries in systems making less complete use of situation. In [35] In the first section of this paper they display a common and shown overview to wavelet analysis. Superior attention is agreed to the method fleeting waveforms are signified under the several transforms and possible drawbacks are decorated with EEG-related instances. In the second section, they show the application of some of these methods to one of parts of specific concern- automated recognition of epileptic form action in the multichannel EEG.

They describe the wavelet-based levels of a spike recognition system they have industrialized. The enactment of this section of the system has been determined in preliminary clinical study. In [36] Gotman J. et al. presented a technique for the involuntary recognition of seizures in the EEG, individually of the attendance of clinical signs; it is based on the decay of the EEG into basic waves and the recognition of convulsive spurts of recurring action having a frequency between 3 and 20 c/sec.

Humble processes are used to amount the amplitude of waves comparative to the experience, their period and rhythmicity. The assessment of the technique on 24 surface footages and 44 records from intracerebral electrodes showed that it was accomplished of diagnosing several kinds of crisis. False recognitions due to non-epileptic form rhythmic EEG spurts and to objects were rather recurrent but were not a crisis problem because they did not overly grow the EEG drawing and they could be simply recognized by the electroencephalographic.

The program can achieve on-line and concurrently the involuntary recognition of spikes and of crisis in 16 channels. In [37] Enhancements to a current automatic crisis recognition program are escribed. They are designed at captivating into version a larger chronological context and thus refining the specificity of the recognitions. Results were calculated on 293 records from 49 cases, total 5303 h of 16-channel records. They displayed that 24% of the 244 crisis noted were lost by the automatic recognition; in 41% of the crisis, the cases alarm was not forced but the machine made recognitions.

The false recognition percentage was of the order of 1 false recognition per hour of demo. Conclusions are: (1) automatic crisis recognition must be used in combination with a cases alarm button since some crisis, taking unwell defined EEG action, are not distinguished; (2) the automatic recognition acceptable detention of many crisis, clinical and subclinical, for which the fear was not pushed; (3) the low false recognition ratio shows that lower recognition threshold could be recycled, acquiescent improved seizure recognition.

In [38] presented the footage of crisis and barbs is of principal rank in the valuation of epileptic cases. This is not continuously an easy procedure because these proceedings can be infrequent and are typically random. Since the initial days of PC examination of the EEG, scientists have industrialized approaches for the involuntary recognition of spikes and, more newly, of crisis.

The problems are multifaceted because spikes and crisis are not obviously distinct and have very diverse morphologies. Nevertheless, it has been likely to grow involuntary recognition approaches that can be of great help through long-term observing of

epileptic cases. No technique is unconditionally fail-safe and all need human proof, but they save a substantial quantity of time in the clarification of extended records.

Current developments contain recognition of the patterns exact to babies, and the option of cautionary a cases or spectator that a crisis is starting. In [39] relates genetic programming (GP) to achieve involuntary feature selection since inventive feature record with the objective of refining the biased accuracy of a classifier and decreasing the data attributes dimensionality. The tree organization of GP obviously signifies the features, and a new job complete in this study automatically agrees the amount of the attributes selected.

In experimentations on two general epilepsy EEG recognition problems, the recognition accurateness on the GP-depended attributes is important greater than on the innovative attributes. Concurrently, the length of the data attributes for the classifier is abundant lesser than that of the innovative attributes. In [40] Ling Guo proposes a new technique for automatic epilepsy crisis recognition, which uses estimated entropy attributes resulting from multi wavelet transmute and mixes with an (ANN) to recognize the EEG signals concerning the being or absenteeism of crisis.

To the greatest information of the researcher, there happens no like work in the works. A well-known common dataset was used to estimate the presented technique. The high performance acquired for two various recognition problems confirmed the achievement of the technique. In [41] Ling Guo et al. deals with a novel technique of study of EEG signals using comparative wavelet energy, and detection using (ANNs).

In [42] Lima et al. explore the presentation of least squares support vector machines (LS-SVM) to the job of epileptic identification over automatic EEG signal cataloguing. More exactly, the researcher extant a sensitivity enquiry research by earnings of which the accuracy stages showed by normal and least squares SVM classifiers are compared, captivating into explanation the set of the kernel role and of its parameter rate. Outcomes of experimentations showed over several kinds of attributes mined from a standard EEG signal database indication that the sensitivity outlines of the kernel machines are qualitatively alike, together viewing distinguished accuracy in footings of precision and generality. In addition, the accuracy talented by optimally organized

LS-SVM replicas is also quantitatively compared with that achieved by associated methods for the same database.

3.MATERIAL AND METHODS

In this chapter, the proposed method and techniques are explained in detail in order to achieve the objectives set forth for the study. The proposed method is mainly based upon deep learning method.

3.1 Deep Learning

Machine learning methods controls many features of current society: since web hunts to content sifting on social networks to references on e-commerce web pages, and it is progressively current in customer crops like cameras and smartphone. Machine learning methods are used to recognize items in images, record speech into text, competition news objects, posts or crops with users' welfares, and determined pertinent results of search. Progressively, these applications type use of a class of methods called deep learning. Conventional machine learning methods were incomplete in their capacity to procedure natural information in their raw form. For years, creating a pattern classification or machine learning model requisite cautious engineering and substantial area knowledge to design a feature extractor that transmuted the raw data into an appropriate interior demonstration or feature vector from which the learning subsystem, frequently a classifier, might identify or categorize patterns in the input. Representation learning is a group of techniques that lets a machine to be nursed with raw data and to repeatedly determine the representations desired for recognition or classification. Deep learning techniques are representation learning techniques with several stages of representation, gained by comprising humble but nonlinear units that each transmute the representation at one stage into an illustration at an advanced, slightly more abstract stage. With the arrangement of sufficient such alterations, very multifaceted jobs can be learned. For classification tasks, advanced layers of demonstration intensify features of the input that are significant for taste and overpower immaterial variations.

complicated constructions in high dimensional information and is therefore appropriate to various areas of science, government and business. In addition to thrashing archives in image recognition [43] and speech recognition, it has compressed extra machine learning methods at guessing the activity of possible drug molecules, studying subdivision accelerator data, rebuilding brain circuits, and guessing the

belongings of alterations in non-coding DNA on gene expression and disease. Maybe more astonishingly, deep learning has shaped very talented results for several errands in usual language sympathetic, chiefly topic classification, sentimentality study, query answering and language conversion [44]. New learning algorithms and buildings that are presently existence industrialized for deep neural networks will only hurry this growth. Deep learning contains number of techniques such as: deep auto-encoder, convolutional neural network and recurrent neural network.

3.1.1 Recurrent Neural Network (RNN)

The essential feature of a Recurrent Neural Network (RNN) is that the network comprises at minimum one feed-back linking, so the activations can flow round in a loop. That allows the networks to do chronological processing and learn orders, e.g., achieve order classify imitation or chronological connotation/estimate. Recurrent neural network architectures can have several various methods. One public kind contains of a normal Multi-Layer Perceptron (MLP) plus extra loops [45]. These can feat the powerful nonlinear charting competences of the MLP, and also have approximately form of memory. Others have additional static constructions, possibly with each neuron connected to all the others, and may also have stochastic activation functions. For modest constructions and deterministic activation functions, learning can be realized using comparable incline descent events to those foremost to the back propagation algorithm for feed-forward networks. When the activations are stochastic, simulated annealing approaches may be more suitable. The next will appearance at an insufficient of the most significant kinds and features of recurrent networks.

The humblest method of completely recurrent neural network is an MLP with the preceding group of hidden component activations nourishing back into the network lengthways with the inputs. Note that the period has to be discretized, with the activations efficient at separately time step. The time gage might agree to the operation of actual neurons, or for artificial methods any time stage size suitable for the assumed problem can be used. A postponement component wants to be presented to hold activations pending they are treated at the following time step see in **Figure 3.1.** The input represented the features that selected by the user and the output is the predication of the network. Hidden unit is the important part for learning in the RNN, which

contain time delay feature that's important to learn sensitive features in time series problems such as signal processing and speech recognition problems.

Figure 3.1: Simplest form of fully recurrent neural network [46]

Then one can contemplate about recurrent networks in footings of their possessions as dynamical schemes, it is usual to ask about their constancy, controllability and observability: Constancy anxieties the boundedness terminated time of the network productions, and the reply of the network productions to minor vicissitudes. Controllability is worried with whether it is likely to switch the dynamic conduct. A recurrent neural network is supposed to be manageable if an early national is steerable to any wanted national inside a limited amount of time steps. Observability is worried with whether it is likely to detect the consequences of the switch practical [47]. A recurrent network is supposed to be noticeable if the state of the network can be strongminded from a limited set of input/output capacities. A hard action of these subjects is way outside the scope of this unit!

Shortening the outspread network to fair one-time step decreases it to a so-called Humble Recurrent Network (which is also generally recognized as an Elman network):

In Figure 3.2 each group of weights now only seems merely once, so it is likely to relate the incline descent method using the normal backpropagation algorithm

somewhat than full Backpropagation through time (BPTT) Backpropagation Through Time, or BPTT, is the training algorithm used to update weights in recurrent neural networks like LSTMs. This income that the mistake signal will not get spread spinal very far, and it will be problematic for the network to learn how to use data from far back in time. In repetition, this estimate shows to be too countless for numerous applied applications.

Figure 3.2: Simple Recurrent Networks [48]

3.1.2 Convolutional Neural Network (CNN)

A particular type of such a deep neural network is the convolutional network, which is usually referred to as CNN or ConvNet. It's a deep neural network. Recall that feed-forward neural networks are also named multi-layer perceptron's (MLPs), which are the typical deep learning replicas. The replicas are called "feed-forward" because data currents correct through the perfect. There are no feedback connections in which outputs of the model are fed back into itself.

CNNs exactly are enthused by the biological graphic cortex. The cortex has minor areas of cells that are subtle to the exact parts of the visual field. In this experimentation, the scientists displayed that approximately separate neurons in the brain triggered or fired only in the attendance of limits of a specific location like perpendicular or straight edges. For instance, some neurons ablaze when bare to perpendicular flanks and some when presented a flat edge. Hubel and Wiesel found

that totally of these neurons were healthy well-ordered in a columnar style and that composed they were talented to crop visual insight. This idea of particular mechanisms intimate of a structure having exact jobs is one that machineries use as well and one that you can also find back in CNNs [49].

CNN have been one of the greatest powerful novelties in the arena of computer vision. They have achieved a lot healthier than traditional computer vision and have shaped state-of-the-art outcomes. These neural networks have established to be fruitful in numerous various real-life case studies and applications [50],

- Image recognition, object classification, face recognition;
- Self-driving cars that leverage CNN based vision systems;
- Classification of crystal structure using a CNN;
- And others, of course!

As shown in **Figure 3.3** input layer mean the data that will be analyzed by the network, convolutional layer is used to extracted features and used share Wight technique for optimal using the memory. Pooling layer used to reduce the dimension of the data. The last stage is the fully connected layer which is used to classify the linear extracted features in to labels.

Figure 3.3: Convolutional Neural Network [51]

3.1.3 Deep Neural Network

The greedy layer wise method for feature extraction and dimension reduction a deep auto encoder neural network works by training each layer alone. In this study, we will work on how auto encoders can be stacked in a greedy layer wise method for pertaining (initializing) the weights of a deep network.

A stacked auto encoder is a type of deep neural network consisting of multiple layers of sparse auto encoders in which the outputs of each auto encoder is connected to the inputs of the successive layer. Officially, consider a stacked auto encoder with n layers. Using record from the auto encoders part, let W(k,1), W(k,2), b(k,1),b(k,2) indicate the parameters W(1),W(2),b(1),b(2) for kth auto encoder. Then the encoding step for the stacked auto encoder is given by running the encoding step of each layer in forward order:

The step of decoding is given by implementing the decoding stack of each auto encoder in inverse order:

The input data of interest is contained within a(n), which is the activation of the deep layer of hidden units. The produce vector gives us a good representation of the input in terms of higher-order features.

3.2 Power Spectral Density (PSD)

In this part briefly we will explain how the strength of a signal is divided in the frequency domain, proportional to the strengths of any signals in the environment, is focal to the design of all Linear time-invariant filter intentional to quell or selection the signal [52]. This good when signals are deterministic, and it convert out to be simply as right in the status of random signals. For example, if a metrical waveform is audio signals with collective disorder signals, its necessary generate a low pass Linear time-invariant filter for extracting the audio and repress the disorder signals [53,54]. Power spectral density function present the force of the energy in the signal as a function of frequency. The unit of Power spectral density function is energy (difference) for each frequency(width) and can gain energy in a certain frequency domain by merge Power spectral density in that frequency domain.

3.3 Proposed Method

In the first stage, PSD is applied to reduce the dimension of the input signal and extract the important features. This stage is very important to reduce the computation time and improve the classification results. PSD is used with a number of classifiers and has shown satisfactory results in previous studies.

As mentioned earlier, in this thesis PSD is used with deep auto-encoder to classify the EEG signals indicative of epilepsy. The PSD value is calculated for each of the 4096 vectors. The PSD output, later, becomes the input to the deep auto-encoder in the next step.

Next, the first auto-encoder converts 4096 inputs to 1924 outputs, thus reducing the dimension of the data in half. The output of the first auto-encoder becomes input to the second auto-encoder. Then, the second auto-encoder converts the 1924 inputs to just 100 as shown in Equations.

A stacked auto encoder (SAE) is a neural network generated from multiple layers of sparse auto-encoders once the outputs of each layer (auto-encoder) turn into inputs for the next layer. In the mathematical sense, in order to represent the SAE, the number of auto-encoders is represented as n layers . Let $W^{(k,1)}$, $W^{(k,2)}$, $b^{(k,1)}$, $b^{(k,2)}$ indicate the parameters $W^{(1)}$, $W^{(2)}$,$b^{(1)}$,$b^{(2)}$ for the kth auto-encoder. The encoding level for the SAE is given by implementing the encoding stage of each layer in forward order:

$$a^{(t)} = f\,(z^{(l)}) \tag{3.1}$$

$$z^{(t+1)} = W^{(l,1)} \quad a^{(l)} \quad + b^{(l+1)} \tag{3.2}$$

The decoding level is specific by executing the decoding with the stacked auto-encoder in inverse order:

$$a(n+l) = f\,(z(n+l)) \tag{3.3}$$

$$z(n+l+1) = W(n-l,2) \quad a(n+l) \quad + b(n-l,2) \tag{3.4}$$

The input of the benefit is included within a(n), which is the activation of the deepest layer of the hidden parts. This array provides a good example of of the information in terms of high-level features. By passing a(n) to a SOFTMAX function, the features produced by the stacked auto encoder can be used for classification problems to deal with problems related to classification.

Epileptic seizures will be detected in this way using our new method. Epileptic seizure is a very sensitive medical problem and each case consists of 4096 features for use by any classifier directly to detect the presence of seizures; yet, this process will naturally take a long time. Our idea focuses on using a given number of auto encoders to represent data in a new form, extract important features and reduce the size of the data. The structure of our network will be in the following form: 4096-1924-100-2. See Figure 3.1.

As stated earlier, the input data consists of 4096 features, two hidden layers are used to reduce the dimensions of the data and feature extraction as 1924 and 100, respectively. The last layer is the output layer, which consists of 2 perceptron's that represent the types of epileptic seizure. See Figure 3.2.

Fig 3.1: Deep neural network based on PSD for Epileptic seizure

Given this information\ the main idea is to train each auto-encoder as unsupervised learning features\ thereby reducing the size of the data, and to train SoftMax as supervised learning by using labeled data. The last stage is the more dramatic stage, which consists of stacking the two unsupervised levels with the last classifier trained as supervised learning, and to train all of them together using the labeled data in a supervised form.

4.RESULTS

Deep auto encoder based on power spectral density is used to detect epileptic seizures with satisfactory results when compared to previous researches. The proposed framework is trained and tested using the dataset proposed in [55] which consists of two groups A, and E. A represents the healthy cases and E represents the epileptic activity. Each case consists of 4096 data. Using our method in the first stage, we calculate the PSD values for the case with 4096 points. The size of the features according to the behavior of the input data. However, the extracted features from PSD calculation stage are shown in the Fig.2, making them the input to the first auto encoder, which converts the 4096 inputs to 1924. This operation reduces the dimension of the data to half, implying savings on time. Then, the output of the first auto encoder becomes the input to the second auto encoder. Here, the second auto encoder converts the 1924 input to the 100 outputs.

Fig 2: PSD of Train Data for Epileptic seizure (Normal and Abnormal)

Fig 3: PSD of the Test Data for Epileptic seizure (Normal and Abnormal)

SoftMax is used to classify the extracted features from the last auto-encoder to its labels to determine whether the signs of epileptic seizures are present. The last procedure stacks the auto encoders and SoftMax layers and trains them together in a supervised fashion. The results of the proposed framework are presented in Figure 4.

Fig.4 Results of deep Auto encoder based on PSD

Then, the true positive result is 47%, true negative is 49% and the total accuracy is the total of two previous results which mean 96%.

The false posetive and false true results are 1% and 3% repectivly. Moreover the sensitivity and specificity are 94% and 98%.

Furthermore, positive predictive value is 97.9 and negative predictive value is 94.2 these parameters have great role to improve the method performance beside the total accuracy.

Table I: Comparison of Results with Previous Works

Referance	Methods	Accuracy %
Srinivasan et al. [56]	Time–frequency domain feature-Recurrent neural network	99.60
Subasi and Ercelebi [57]	WT+ANN	92.00
Subasi et al. [58]	Discrete WT-Mixture of expert model	94.5
Kannathal et al. [59]	Entropy measures-ANFIS	92.22
Tzallas et al. [60]	Time–frequency analysis—ANN	100
Polat et al. [61]	Fast Fourier transform-DT	98.72
Acharya et al. [63]	WPD-PCA-GMM	99.00
Acharya et al [64]	Entropies + HOS +Higuchi FD+ Hurst exponent+ FC	99.70
Proposed Framework	**power Spectral Density+ Deep autoencoders+ SoftMax**	**96**

5. DISCUSSION OF COMMON ATTRIBUTES AND CONCLUSION

5.1 Discussion

In Table 1, the results obtained in the field of epilepsy seizure detection are compared with a number of other most noted works. Accordingly, some studies have applied common methods in the field of automatic detection of epilepsy serious. In comparison, the method presented in this thesis surpasses in terms of the outcomes those proposed in [57], [58], and [59].

Yet, there are other studies with better results compared to those obtained herein, namely [56], [60], [61], [62], [63], and [64].

However, there are two advantages to our system when compared with all studies presented in the Table 1; they applied only the conventional feature extraction and classification methods, which require a long time to process. The second issue is that our method is trained using only 50% of the data; whereas, the others had to use 60% and more.

5.2 Conclusion

In this thesis, deep auto-encoders used based on power spectral density (PSD) method to detect incidents of epilepsy in the EEG records of patients. The aim of the proposed method is to calculate the power spectral density for each input signal and to use deep sparse auto-encoder to classify the signals into normal and abnormal groups. The proposed method presented satisfactory and new results when compared with previous research. The reason to employ PSD before deep sparse auto-encoder is to extract features from the input data and speed up the processing of the data.

The output of PSD becomes, in this way, the input to the two subsequent auto-encoders, which reduce the features number step by step and extract high-level features. Then, SoftMx is used to classify these extracted features into two labels, and the two auto-encoders are combined with SoftMax and trained in the form of supervised learning to increase the performance of the detection. The proposed method presented 96% accuracy in few time when compared with other methods shown in Table 1.

5.3 Limitation of Study

The first limitation is that is short time to prepare a master's thesis once a topic as new as deep learning is concerned, thus the absence of other studies in the field. Furthermore, some techniques require special hardware, such as CNN and LSTM, which need hardware's such as CUDA, an expensive and rare item. Moreover, this problem can be solved by using CNN and LSTM in the feature because they can present more satisfactory results in this type of problems.

5.4 Future Works

As future work, the proposed framework can be used in other various fields, such as computer network security, speech recognition and video classification. Furthermore, the framework can be developed using algorithms to optimize the performance of auto encoders or more effective feature extraction techniques instead of power spectral density.

Furthermore, power spectral density can be used with different deep learning techniques such as CNN, LSTM, RNN.

5. References

[1] B. Chandra and R. K. Sharma, "Deep learning with adaptive learning rate using laplacian score," *Expert Syst. Appl.*, vol. 63, pp. 1–7, 2016.

[2] G. Urban *et al.*, "Do Deep Convolutional Nets Really Need to be Deep and Convolutional?," 2016.

[3] "3. Đorđe T. Grozdića.pdf." .

[4] K. G. Lore, A. Akintayo, and S. Sarkar, "LLNet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognit.*, vol. 61, pp. 650–662, 2017.

[5] K. Sun, J. Zhang, C. Zhang, and J. Hu, "Generalized extreme learning machine autoencoder and a new deep neural network," *Neurocomputing*, vol. 230, no. November 2016, pp. 374–381, 2017.

[6] Y. Xiong and R. Zuo, "Recognition of geochemical anomalies using a deep autoencoder network," *Comput. Geosci.*, vol. 86, pp. 75–82, 2016.

[7] L. D. Burgoon, "Autoencoder Predicting Estrogenic Chemical Substances (APECS): An improved approach for screening potentially estrogenic chemicals using in vitro assays and deep learning," *Comput. Toxicol.*, vol. 2, pp. 45–49, 2017.

[8] C. Hong, J. Yu, J. Wan, D. Tao, and M. Wang, "Multimodal Deep Autoencoder for Human Pose Recovery," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5659–5670, 2015.

[9] T. H. Song, V. Sanchez, H. Eidaly, and N. M. Rajpoot, "Hybrid deep autoencoder with Curvature Gaussian for detection of various types of cells in bone marrow trephine biopsy images," *Proc. - Int. Symp. Biomed. Imaging*, pp. 1040–1043, 2017.

[10] Y. Suzuki and T. Ozaki, "Stacked denoising autoencoder-based deep collaborative filtering using the change of similarity," *Proc. - 31st IEEE Int. Conf. Adv. Inf. Netw. Appl. Work. WAINA 2017*, pp. 498–502, 2017.

[11] Y. D. Zhang, X. X. Hou, Y. D. Lv, H. Chen, Y. Zhang, and S. H. Wang, "Sparse autoencoder based deep neural network for voxelwise detection of cerebral microbleed," *Proc. Int. Conf. Parallel Distrib. Syst. - ICPADS*, vol. 164, pp. 1229–1232, 2017.

[12] M. Peker, "A new approach for automatic sleep scoring: Combining Taguchi based complex-valued neural network and complex wavelet transform," *Comput. Methods Programs Biomed.*, vol. 129, pp. 203–216, 2016.

[13] M. Peker, B. Şen, and P. Y. Kumru, "An efficient solving of the traveling salesman problem: The ant colony system having parameters optimized by the Taguchi method," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 21, no. SUPPL. 1, pp. 2015–2036, 2013.

[14] M. Nandhini, B. Suchithra, R. Saravana¬thamizhan, and D. G. Prakash, "Optimization of parameters for dye removal by electro- -oxidation using Taguchi Design," *J. Electrochem. Sci. Eng.*, vol. 4, no. 4, pp. 227–234, 2014.

[15] L. Ivanović, B. Stojanović, J. Blagojević, G. Bogdanović, and A. Marinković, "Analysis of the flow rate and the volumetric efficiency of the trochoidal pump by application of taguchi method," *Teh. Vjesn.*, vol. 24, pp. 265–270, 2017.

[16] N. Hoque, H. Kashyap, and D. K. Bhattacharyya, "Real-time DDoS attack detection using FPGA," *Comput. Commun.*, vol. 110, pp. 48–58, 2017.

[17] B. B. Gupta, "Predicting number of zombies in DDoS attacks using pace regression model," *J. Comput. Inf. Technol.*, vol. 20, no. 1, pp. 33–39, 2012.

[18] M. Alkasassbeh, A. B. A. Hassanat, and G. Al-naymat, "Detecting Distributed Denial of Service Attacks Using Data Mining Techniques," vol. 7, no. 1, pp. 436–445, 2016.

[19] Y. Ju, J. Guo, and S. Liu, "A Deep Learning Method Combined Sparse Autoencoder with SVM," *Proc. - 2015 Int. Conf. Cyber-Enabled Distrib. Comput. Knowl. Discov. CyberC 2015*, pp. 257–260, 2015.

[20] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," *2015 Mil. Commun. Inf. Syst. Conf.*, pp. 1–6, 2015.

[21] S. M. Hosseini Bamakan, H. Wang, and Y. Shi, "Ramp loss K-Support Vector Classification-Regression; a robust and sparse multi-class approach to the intrusion detection problem,"

Knowledge-Based Syst., vol. 126, pp. 113–126, 2017.
[22] C. Khammassi and S. Krichen, "A GA-LR wrapper approach for feature selection in network intrusion detection," *Comput. Secur.*, vol. 70, pp. 255–277, 2017.
[23] S. Yücel, P. Terzioğlu, and D. Özçimen, "World â€™ s largest Science , Technology & Medicine Open Access book publisher c," *RFID Technol. Secur. Vulnerabilities, Countermeas.*, 2016.
[24] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, "Seizure prediction: The long and winding road," *Brain*, vol. 130, no. 2, pp. 314–333, 2007.
[25] P. Evidence, B. Eb, and P. Neprovociranom, "EVIDENCE BASE (EB) APROACH TO THE FIRST UNPROVOKED."
[26] C. J. James and B. E. Eng, "Detection of epileptiform activity in the electroencephalogram using artificial neural networks," no. February 1997, 1997.
[27] A. M. Karim, F. V. Çelebi, and A. S. Mohammed, "Software Development for Blood Disease Expert System," Lect. Notes Softw. Eng., vol. 4, no. 3, pp. 179–183, 2016.
[28] B. Chandra and R. K. Sharma, "Deep learning with adaptive learning rate using laplacian score," *Expert Syst. Appl.*, vol. 63, pp. 1–7, 2016.
[29] G. Urban *et al.*, "Do Deep Convolutional Nets Really Need to be Deep and Convolutional?," 2016.
[30] K. G. Lore, A. Akintayo, and S. Sarkar, "LLNet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognit.*, vol. 61, pp. 650–662, 2017.
[32] K. Sun, J. Zhang, C. Zhang, and J. Hu, "Generalized extreme learning machine autoencoder and a new deep neural network," *Neurocomputing*, vol. 230, no. November 2016, pp. 374–381, 2017.
[33] Y. Xiong and R. Zuo, "Recognition of geochemical anomalies using a deep autoencoder network," *Comput. Geosci.*, vol. 86, pp. 75–82, 2016.
[34] L. D. Burgoon, "Autoencoder Predicting Estrogenic Chemical Substances (APECS): An improved approach for screening potentially estrogenic chemicals using in vitro assays and deep learning," *Comput. Toxicol.*, vol. 2, pp. 45–49, 2017.
[35] C. Hong, J. Yu, J. Wan, D. Tao, and M. Wang, "Multimodal Deep Autoencoder for Human Pose Recovery," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5659–5670, 2015.
[36] T. H. Song, V. Sanchez, H. Eidaly, and N. M. Rajpoot, "Hybrid deep autoencoder with Curvature Gaussian for detection of various types of cells in bone marrow trephine biopsy images," *Proc. - Int. Symp. Biomed. Imaging*, pp. 1040–1043, 2017.
[37] Y. Suzuki and T. Ozaki, "Stacked denoising autoencoder-based deep collaborative filtering using the change of similarity," *Proc. - 31st IEEE Int. Conf. Adv. Inf. Netw. Appl. Work. WAINA 2017*, pp. 498–502, 2017.
[38] Y. D. Zhang, X. X. Hou, Y. D. Lv, H. Chen, Y. Zhang, and S. H. Wang, "Sparse autoencoder based deep neural network for voxelwise detection of cerebral microbleed," *Proc. Int. Conf. Parallel Distrib. Syst. - ICPADS*, vol. 164, pp. 1229–1232, 2017.
[39] M. Peker, "A new approach for automatic sleep scoring: Combining Taguchi based complex-valued neural network and complex wavelet transform," *Comput. Methods Programs Biomed.*, vol. 129, pp. 203–216, 2016.
[40] M. Peker, B. Şen, and P. Y. Kumru, "An efficient solving of the traveling salesman problem: The ant colony system having parameters optimized by the Taguchi method," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 21, no. SUPPL. 1, pp. 2015–2036, 2013.
[41] M. Nandhini, B. Suchithra, R. Saravana¬thamizhan, and D. G. Prakash, "Optimization of parameters for dye removal by electro- -oxidation using Taguchi Design," *J. Electrochem. Sci. Eng.*, vol. 4, no. 4, pp. 227–234, 2014.
[42] L. Ivanović, B. Stojanović, J. Blagojević, G. Bogdanović, and A. Marinković, "Analysis of the flow rate and the volumetric efficiency of the trochoidal pump by application of taguchi method," *Teh. Vjesn.*, vol. 24, pp. 265–270, 2017.
[43] N. Hoque, H. Kashyap, and D. K. Bhattacharyya, "Real-time DDoS attack detection using FPGA," *Comput. Commun.*, vol. 110, pp. 48–58, 2017.
[44] B. B. Gupta, "Predicting number of zombies in DDoS attacks using pace regression model," *J. Comput. Inf. Technol.*, vol. 20, no. 1, pp. 33–39, 2012.
[45] M. Alkasassbeh, A. B. A. Hassanat, and G. Al-naymat, "Detecting Distributed Denial of Service Attacks Using Data Mining Techniques," vol. 7, no. 1, pp. 436–445, 2016.
[46] Y. Ju, J. Guo, and S. Liu, "A Deep Learning Method Combined Sparse Autoencoder with SVM," *Proc. - 2015 Int. Conf. Cyber-Enabled Distrib. Comput. Knowl. Discov. CyberC 2015*,

pp. 257–260, 2015.
[47] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," *2015 Mil. Commun. Inf. Syst. Conf.*, pp. 1–6, 2015.
[48] S. M. Hosseini Bamakan, H. Wang, and Y. Shi, "Ramp loss K-Support Vector Classification-Regression; a robust and sparse multi-class approach to the intrusion detection problem," *Knowledge-Based Syst.*, vol. 126, pp. 113–126, 2017.
[49] C. Khammassi and S. Krichen, "A GA-LR wrapper approach for feature selection in network intrusion detection," *Comput. Secur.*, vol. 70, pp. 255–277, 2017.
[50] S. Yücel, P. Terzioğlu, and D. Özçimen, "World â€™ s largest Science , Technology & Medicine Open Access book publisher c," *RFID Technol. Secur. Vulnerabilities, Countermeas.*, 2016.
[51] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, "Seizure prediction: The long and winding road," *Brain*, vol. 130, no. 2, pp. 314–333, 2007.
[52] P. Evidence, B. Eb, and P. Neprovociranom, "EVIDENCE BASE (EB) APROACH TO THE FIRST UNPROVOKED."
[53] C. J. James and B. E. Eng, "Detection of epileptiform activity in the electroencephalogram using artificial neural networks," no. February 1997, 1997.
[54] Y. Xiong and R. Zuo, "Recognition of geochemical anomalies using a deep autoencoder network," *Comput. Geosci.*, vol. 86, pp. 75–82, 2016.
[55] L. D. Burgoon, "Autoencoder Predicting Estrogenic Chemical Substances (APECS): An improved approach for screening potentially estrogenic chemicals using in vitro assays and deep learning," *Comput. Toxicol.*, vol. 2, pp. 45–49, 2017.
[56] V. Srinivasan, C. Eswaran, and A. N. Sriraam, "Artificial neural network based epileptic detection using time-domain and frequency-domain features," J. Med. Syst., vol. 29, no. 6, pp. 647–660, 2005.
[57] A. Subasi and E. Erçelebi, "Classification of EEG signals using neural network and logistic regression," Comput. Methods Programs Biomed., vol. 78, no. 2, pp. 87–99, 2005.
[58] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," Expert Syst. Appl., vol. 32, no. 4, pp. 1084–1093, 2007.
[59] N. Kannathal, M. L. Choo, U. R. Acharya, and P. K. Sadasivan, "Entropies for detection of epilepsy in EEG," Comput. Methods Programs Biomed., vol. 80, no. 3, pp. 187–194, 2005.
[60] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Automatic seizure detection based on time-frequency analysis and artificial neural networks," Comput. Intell. Neurosci., vol. 2007, 2007.
[61] K. Polat and S. Güneş, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform," Appl. Math. Comput., vol. 187, no. 2, pp. 1017–1026, 2007.
[62] U. Rajendra Acharya, S. Vinitha Sree, A. P. C. Alvin, and J. S. Suri, "Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework," Expert Syst. Appl., vol. 39, no. 10, pp. 9072–9078, 2012.
[63] U. R. ACHARYA, S. V. SREE, P. C. A. ANG, R. YANTI, and J. S. SURI, "Application of Non-Linear and Wavelet Based Features for the Automated Identification of Epileptic Eeg Signals," Int. J. Neural Syst., vol. 22, no. 2, p. 1250002, 2012.
[64] M. Peker, B. Sen, and D. Delen, "A Novel Method for Automated Diagnosis of Epilepsy using Complex- alued Classifiers.," IEEE J. Biomed. Heal. informatics, vol. 2194, no. c, pp. 1–11, 2015.
[65] A. M. Karim, M. S. Güzel, M. R. Tolun, H. Kaya, and F. V Çelebi, "A New Generalized Deep Learning Framework Combining Sparse Auto-encoder and Taguchi Method for Novel Data Classification and Processing," pp. 1–22
[66] A. M. Karim, F. V. Çelebi, and A. S. Mohammed, "Software Development for Blood Disease Expert System," Lect. Notes Softw. Eng., vol. 4, no. 3, pp. 179–183, 2016.