Study on Machine Learning Algorithms for the Epilepsy Automatic Diagnosis

癲癇の自動診断における機械学習法に関する研究

by

XUYANG ZHAO

Submitted in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

Supervised by Prof. JIANTING CAO

Department of Electronic Engineering
Graduate School of Engineering
Saitama Institute of Technology

Fukaya, Saitama

March, 2020

Acknowledgments

I would like to express my special appreciation and thanks to my advisor professor Jianting Cao, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a researcher. Your advice on both research as well as on my career have been priceless. I would like to thank doctor Qibin Zhao, the unit leader at Tensor Learning Unit, RIKEN Center for Advanced Intelligence Project, to provided me with a research platform and taught me how to become a researcher. I would also like to thank professor Toshihisa Tanaka, Professor of Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology. Let me participate in research projects and learned a lot in the project. I would also like to thank my committee members, professor Hirokazu Yoshizawa, professor Takaharu Yamazaki and professor Tomomi Hashimoto for serving as my committee members. I also want to thank you for letting my defense be an enjoyable moment, and for your brilliant comments and suggestions, thanks to you. A special thanks to my family. Words cannot express how grateful I am to my parents for all of the sacrifices that you've made on my behalf. Your prayer for me was what sustained me thus far. I would also like to thank all of my friends who supported me in writing and incented me to strive towards my goal.

Abstract

Epilepsy is a chronic disorder of the brain caused by excessive discharge of brain cells. Currently, doctors are manually diagnosing by visual judgment based on long-term intracranial electroencephalogram (iEEG) data. This diagnostic procedure is very time consuming and relies on experience. In order to reduce the workload of the doctor, an automatic diagnosis system with high accuracy is required.

In recent years, machine learning methods are often applied to medical diagnosis field. In this thesis, we proposed several methods using conventional machine learning and deep learning. Conventional machine learning methods use filters, entropy and short time Fourier transform (STFT) to extract features. Next, based on the labeled data, the support vector machine (SVM), neural network, and convolutional neural network (CNN) classification models were used for focal localization problems and achieve high performance. However, when these methods are applied to practical clinical patient data, it is difficult to acquire a large amount of iEEG data with high quality labels.

Aim at this problem, this study uses a weakly supervised learning method (PU learning) to learn a binary classifier with a small amount of labeled data and a large amount of unlabeled data. The labeling work can be greatly reduced. In addition, a data augmentation method is used based on the discrete cosine transform (DCT) to generate a large amount of artificial data. Combine the raw data and artificial data as a training data set, the performance

of the model is improved. Further, an epilepsy diagnosis and treatment system was proposed and effectively improving the diagnosis and treatment process.

The structure of this thesis is as follows. Chapter 1 is an introduction and describes the purpose and background of the research. Chapter 2 introduces the principles, statistical features, and datasets of the scalp and intracranial EEG. Chapter 3 describes the feature extraction method and its experimental results. Chapter 4 describes the deep learning method and its experimental results. Chapter 5 introduces the weakly supervised learning method and explains its mathematical principle and experimental results. Chapter 6 describes the data augmentation method and its experimental results. Chapter 7 Summary of research and future issues are described.

Abstract (和文)

癲癇には、脳細胞の過剰な放電によって引き起こされる脳の慢性障害病気である。現在、医師は長時間記録された頭蓋内脳波（iEEG）データに基き、手動的に目視判断による癲癇を診断している。この診断法は非常に時間がかかることと経験に依存している。医師の作業負荷を軽減するために、高精度癲癇の自動診断システムが必要である。

近年、機械学習法は医学分野の診断に適用されることが多い。本論文では、従来の機械学習法と深層学習法を使用し、いくつかの手法を提案した。従来の機械学習方法では、フィルタ、エントロピーおよび短時間フーリエ変換（STFT）を使用し、特徴を抽出する。またウェーブレット変換（WT）、経験モード分解（EMD）などにより、適切な特徴が抽出される。次にラベル付きデータに基づいて、サポートベクターマシン（SVM）、ニューラルネットワーク、畳み込みニューラルネットワーク（CNN）の分類モデルを使用し、癲癇焦点の局所化の高いパフォーマンスの実現ができた。これらの方法を患者実測データに適用する場合には、大量の高品質なラベルが付いたデータを取得することは難しい。

このため、本研究では、弱教師あり学習法を利用することで、少量のラベル付きデータと大量のラベルなしデータでバイナリ分類器を学習でき、それによれば、医師のラベル作業が大幅に削減されることができる。また、離散余弦変換に基づいてデータ増強法を使用し、大量的な人工データを生成し、実データと人工

データをトレーニングデータセットとして組み合わせる。モデルのパフォーマンスを向上した。さらに、癲癇の補助治療を支援するため手法を提案し、診断と治療プロセスを効果的に改善した。

本論文の構成は下記の通りである。第１章は概論で、研究の目的と背景を述べている。第２章は頭皮脳波と頭蓋内脳波の原理、統計学特徴、データセットについて紹介する。第３章は特徴抽出法による実験結果についてを述べる。第４章は深層学習法とその実験結果を述べる。第５章は弱教師ありの学習法を紹介し、その数学原理と実験結果について説明する。第６章はデータ増強法と実験結果を述べている。第７章研究のまとめと今後の課題について述べる。

Contents

List of Figures

List of Tables

1. Introduction

1.1 Epilepsy

Epilepsy is a chronic neurological disorder of the brain which caused by excessive electrical discharges that affects people globally, approximately 65 million people have epilepsy (Epilepsy Foundation of America) [1]. In terms of pediatric patients, 30% of them will face intellectual disability and neurological disorders problems. On the other hand, every year, 50-100 traffic accidents caused by epilepsy seizures are happening in Japan. Some of the epileptic patients are cured or control the frequency of seizures using antiepileptic drugs, but for drug-resistant patients, medication is ineffect!ive and they still face life problems caused by epilepsy. The International League Against Epilepsy (ILAE) [2] defined epilepsy by any of the following conditions:

- At least two unprovoked (or reflex) seizures occurring >24 h apart;
- One unprovoked (or reflex) seizure and a probability of further seizures similar to the general recurrence risk (at least 60%) after two unprovoked seizures, occurring over the next 10 years;
- Diagnosis of an epilepsy syndrome.

According to the definition given by the ILAE [3], there are three types of epileptic seizures: focal onset, generalized onset and unknown onset. In this article, we focus on focal onset, which is recurrent seizures together with abnormal discharge in focal brain areas. The patients suffering from focal seizures can be treated with daily medication to

control epileptic seizure frequency, but approximately 30% of focal seizure patients are refractory to medication [4], hence remove the epileptic focus by surgery is considered as a common treatment. Before the surgery, it needs to locate the epileptic focus and it also the most important factor affecting the outcome of surgery. To determine the localization of epileptic focus, physical exam, iEEG, magnetoencephalogram (MEG), functional magnetic resonance imaging (fMRI) and other modalities are usually performed [5–8]. Considering the cost of MEG and the recording condition (can not move freely), MEG is not practical for experimental work. Because the time resolution of fMRI is very low, for some neurodegenerative diseases, the brain activities can not be completely recorded. iEEG is recorded directly from the cortex, which enables clinical experts to analyze the brain activity effective. These advantages make iEEG as a fundamental tool in the detection, diagnosis, and treatment of epilepsy.

While in current clinical practice Fig. 1.1, clinical experts still through visual for judgment and give annotation manually, which is a time-consuming process and subjective process, the diagnostic results from different clinical experts are often not identical, usually clinical experts need to vote on the diagnosis. In addition, the number of clinical experts is far from enough compared to a large number of patients (e.g. only 689 clinical experts in Japan who can do the diagnosis of epilepsy). Therefore, in the diagnosis of epilepsy, there is a very strong demand.

According to the period of iEEG collection, there are two kinds of brain signals, one is interictal signal which is recorded between epileptic seizures, the other one is ictal signal which is recorded during an epileptic seizure, in this article, we use the interictal brain signal data. The channel recorded from the epileptogenic area is named focal channel, and the channel recorded from the non-epileptogenic area is named non-focal channel. During interictal periods, the epileptic spikes are often used for diagnosis of epilepsy, which can classify and localize the epileptic focus. The International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFSECN) [9] give the following

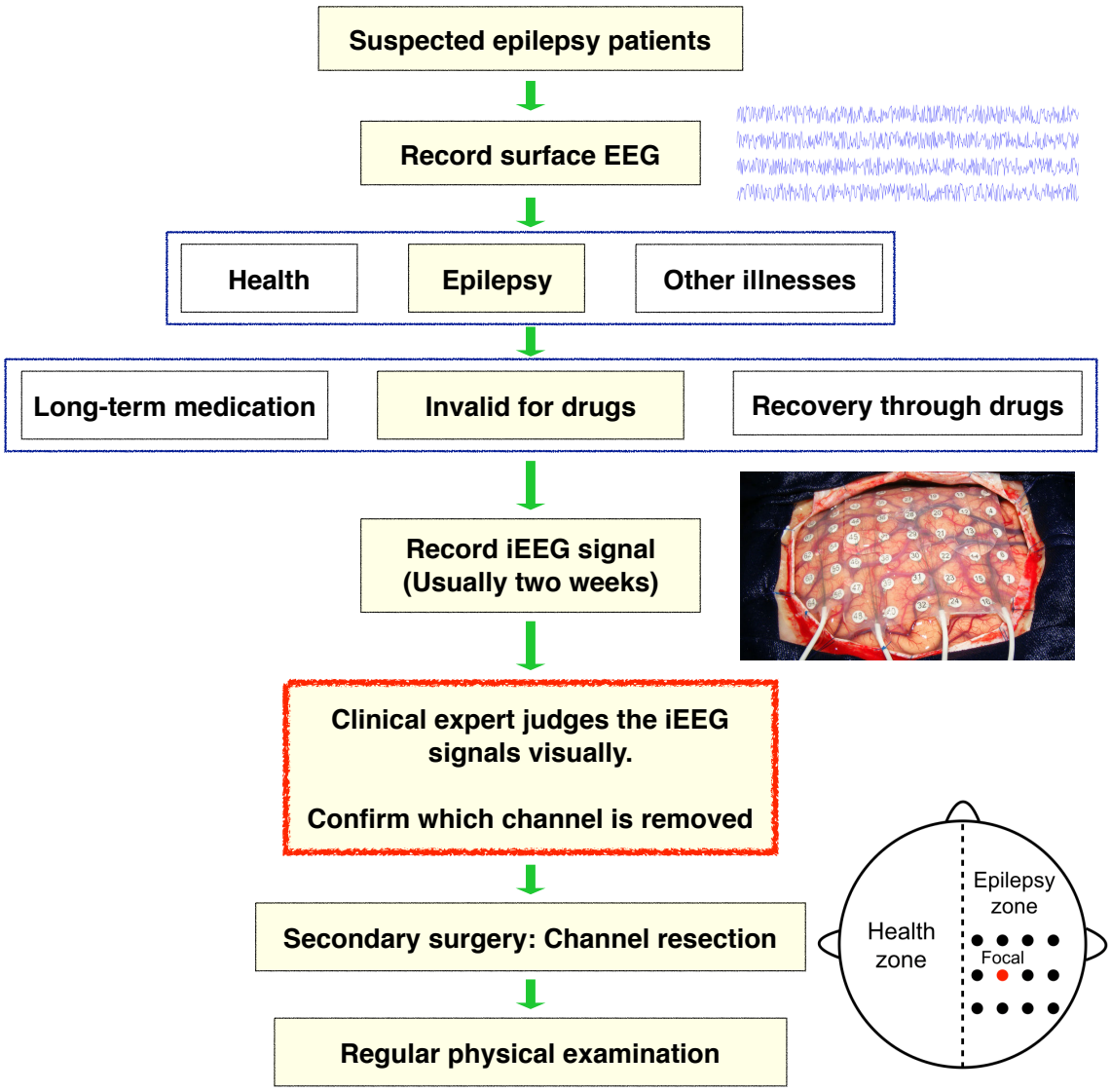

Figure 1.1: Current clinical epilepsy diagnosis procedure.

definitions of spikes:

- Sharp wave: A transient, clearly distinguished from background activity, with pointed peak at a conventional paper speed or time scale and duration of 70 $\pm$ 200 ms, i.e. over 1/4 $\pm$ 1/5 s approximately.
- Spike: A transient, clearly distinguished from background activity, with pointed peak at a conventional paper speed or time scale and a duration from 20 to under 70 ms, i.e. 1/50 $\pm$ 1/15 s, approximately.
- Slow wave: Wave with duration longer than alpha waves, i.e. over 1/8 s.

- Spike-and-slow-wave complex: A pattern consisting of a spike followed by a slow wave.
- Multiple spike complex: A sequence of two or more spikes.
- Polyspikes-and-slow-wave complex: A sequence of two or more spikes associated with one or more slow waves.

1.2 Related Work

In recent years, a variety of methods are proposed for epileptic focus localization, such as template matching [10–15], in this method, first, we need collectted a large amount of sample template waveforms form patients, when we diagnose the new EEG signal, compare with the templates in database to decide whether it is focal or non-focal signal. thereby make automated detection faster than conventional Visual diagnosis. However, this method also has obvious limitations. The samples collected by the database directly affect the diagnosis results. dictionary learning [16], by this method, the author can infer the iEEG from only the scalp EEG by using that dictionary and mapping function. classification [17–21] and some other methods. With the rapid development of artificial neural networks [22], classification methods show a better performance in epileptic focus localization problem. In the classification method, there are two main steps, feature extraction and classification. In the feature extraction step, Discrete wavelet transform (DWT) [23–28], Entropy [29] [30], Fourier transform (FT) [31], Empirical mode decomposition (EMD) [32] and Filter [33] are often used. The resolution of time and frequency is the main advantage of DWT, which makes it especially suitable for the analysis of the non-stationary signals (e.g. iEEG). EMD method decomposes a signal into intrinsic mode functions (IMFs), which is the modulated component of amplitude and frequency. Epilepsy is caused by excessive electrical discharges in brain cells, the entropy method is a measure of the energy which is very suit for feature extraction of epilepsy brain signal.

In the classification step, K-nearest neighbor (KNN), probabilistic neural network (PNN) [17], SVM, FCNN, CNN and other methods are usually be used. In our experiments, several typical supervised learning methods (SVM, FCNN and CNN) are used for classification. Moreover, we introduce the positive unlabeled (PU) learning method to further reduce the workload of clinical experts [34], which solves the label problem very well, and greatly reduces the workload of clinical experts.

1.3 Thesis Contribution

In this thesis, we propose four major methods (feature extraction methods, end-to-end models, weakly supervised learning, data augmentation) 1.2 to solve the practical problems encountered in the clinical diagnosis of epilepsy.

In the traditional feature extraction method, we pay attention to the medical interpretability of the method when selecting the feature extraction method. In order to reduce the manually select feature workload, we propose some end-to-end models which can directly classify the original signal, avoid a lot of feature extraction work. In view of the difficulty in obtaining medical data labels, we introduce weakly supervised learning (PU learning). The annotation work is greatly reduced while ensuring the performance of the model. Because the performance of the method is heavily dependent on the amount of data, we propose a data augmentation method that generates a large amount of artificial data and the performance of the models is improved.

1.4 Thesis Outline

The structure of this thesis is as follows. The first chapter outlines the purpose and background of the research, and the second chapter introduces the principles of intracranial EEG and statistical feature extraction methods. Chapter 3 describes the experimental results of the feature extraction method. Chapter 4 describes the deep learning method and experimental

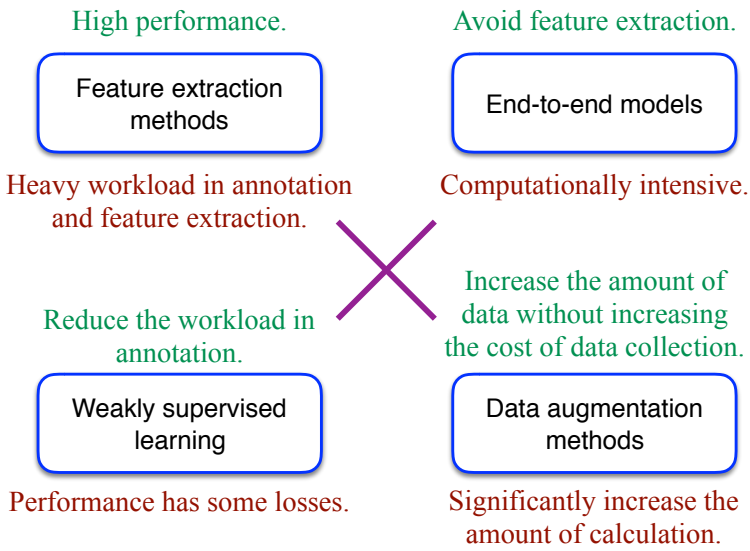

Figure 1.2: Advantages and disadvantages of the four major methods.

results. Chapter 5 introduces a weakly supervised learning method. Chapter 6 describes the data augmentation method. Chapter 7 Summary of research and future issues.

2. Electroencephalogram

EEG examination is a commonly used method for the diagnosis of epilepsy. Compared with other methods (CT, PET, fMRI, etc.), EEG is easy to collect. Epilepsy is a disease caused by abnormal discharge of brain cells, and brain waves can directly record the cell discharge activity of the brain. Currently, EEG is the necessary examination method in the diagnosis of epilepsy. According to the location of EEG acquisition, EEG is divided into scalp EEG and iEEG. We will introduce two kinds of EEG separately.

2.1 Scalp Electroencephalogram

Scalp EEG is a typically noninvasive electrophysiological monitoring method to record the electrical activity of the brain. Scalp EEG is usually used in the diagnose of epilepsy, brain death, for tumors, stroke and other brain disorders. in the thesis, Bonn dataset [35] is used.

2.1.1 Bonn Dataset

Bonn dataset is recorded from The Department of Epileptology in Bonn and is a clinic recognized as an Epilepsy Center in Germany. This dataset includes five sets (A-E), each set containing 100 EEG segments of 23.6 seconds with a sampling rate of 173.61 Hz, all samples are processed with a bandpass filter (0.5–40 [Hz]). Sets A and B are recorded from five healthy volunteers using a standardized electrode placement scheme Fig. 2.1. A set is recorded by volunteers were relaxed in an awake state with eyes open. B set is recorded

by volunteers were relaxed in an awake state with eyes closed. Sets C, D and E record from the patients who had achieved complete seizure control after resection of one of the hippocampal formations. C set recorded from the hippocampal formation of the opposite hemisphere of the brain. D set recorded from within the epileptogenic zone. Sets C and D contained only activity measured during seizure free intervals, E set contained seizure activity. The five sets is shown in Fig. 2.2.

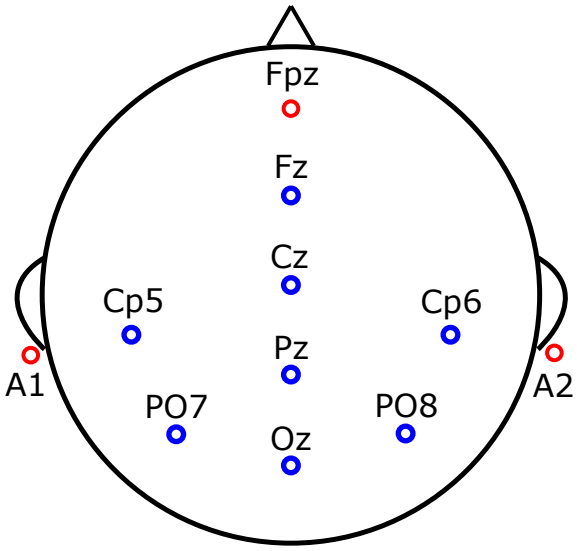

Figure 2.1: International 10–20 system.

2.1.2 Visualization of Bonn Dataset

In order to understand the EEG data more clearly, we use the dimension reduction method to visualize the data, you can find the method introduction and result in web [1]. t-SNE Algorithm is used for dimension reduction which is easily visualized by us humans. The visualize the Bonn data is shown in Fig. 2.3.

[1] https://www.kaggle.com/sabinem/geometry-of-epilepsy-detection/notebook

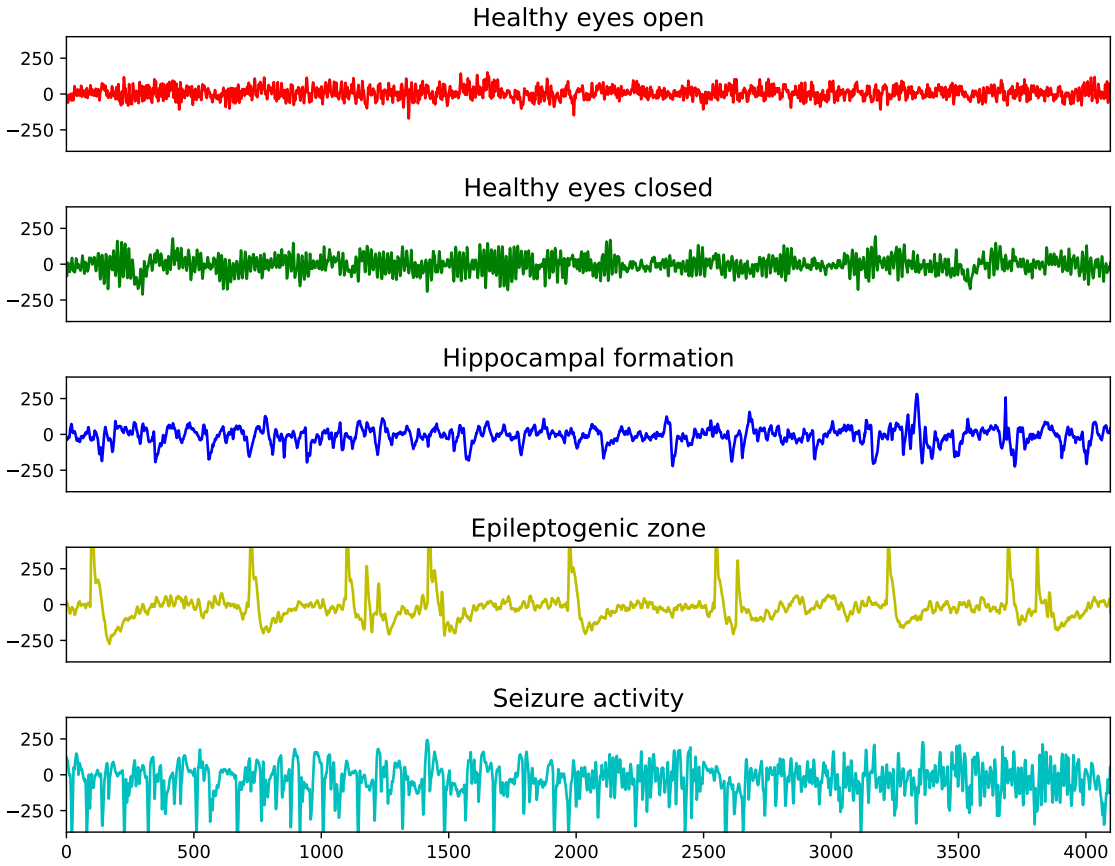

Figure 2.2: Samples of scalp EEG (Bonn Dataset).

2.2 Intracranial Electroencephalogram

iEEG is electrophysiological monitoring that directly records electrical activity from the cerebral cortex. Intracranial electrode placement is shown in Fig. 2.4 (You can find this figure in web [2]) Compare with scalp EEG, iEEG can capture more intracranial information and reduce noise interference. In the thesis, two datasets are used, one is the public dataset of Bern Barcelona Dataset, another is recorded in Juntendo University.

2.2.1 Bern Barcelona Dataset

The dataset [36] includes the iEEG recorded from five patients, who have long standing pharmacoresistant temporal lobe epilepsy and are candidates for epilepsy surgery. iEEG are

[2]https://www.sciencemag.org/news/2019/01/artificial-intelligence-turns-brain-activity-speech

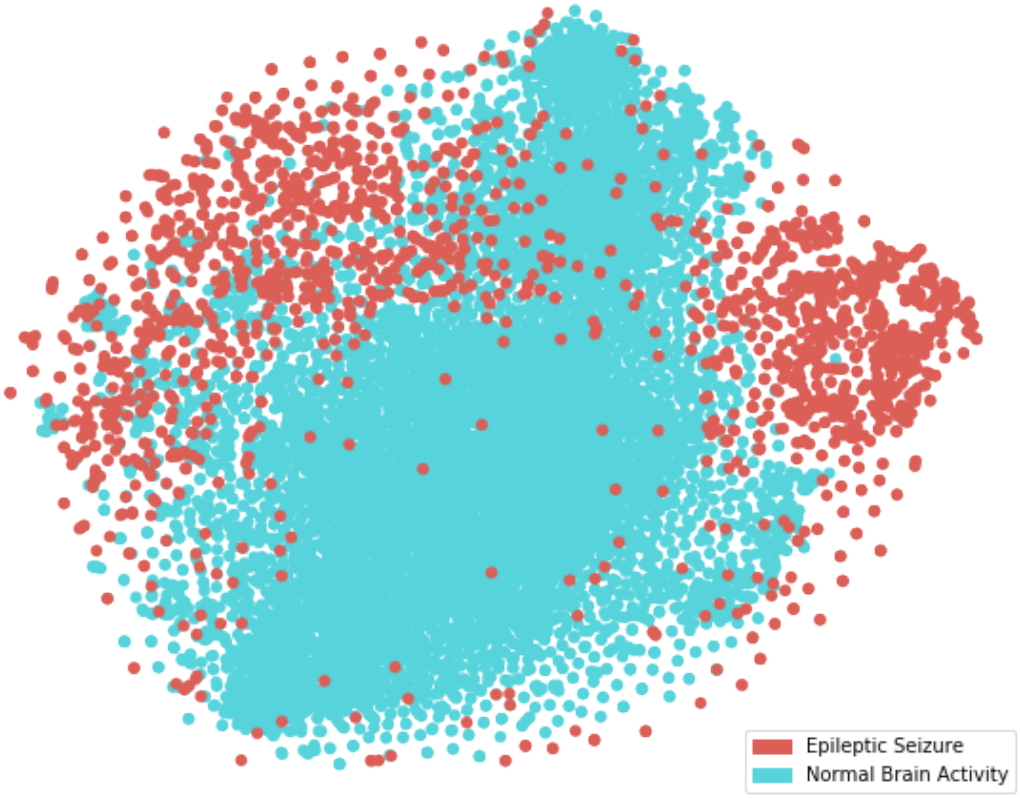

Figure 2.3: Visualization of Bonn dataset (t-SNE).

recorded by the device of AD-TECH (Racine, WI, USA) with the sampling rate of 512 Hz. The channel defined focal channel which detected first ictal iEEG signal changes and judged throw visual by at least two clinical experts, the other channel are defined by non-focal channel. Then, randomly selected 3,750 pairs samples from focal and non-focal channel, respectively, total of 15,000 samples (7,500 focal samples and 7,500 non-focal samples) and each sample is 20 seconds and is processed with the bandpass filter between 0.5 and 150 Hz by using fourth-orders Butterworth filter. An example of focal and non-focal samples are shown in Fig. 2.5.

2.2.2 Juntendo Dataset

The other dataset used in this article is Juntendo dataset, which is recorded at Juntendo University Hospital (Tokyo, Japan). The dataset include four patients who are suffering

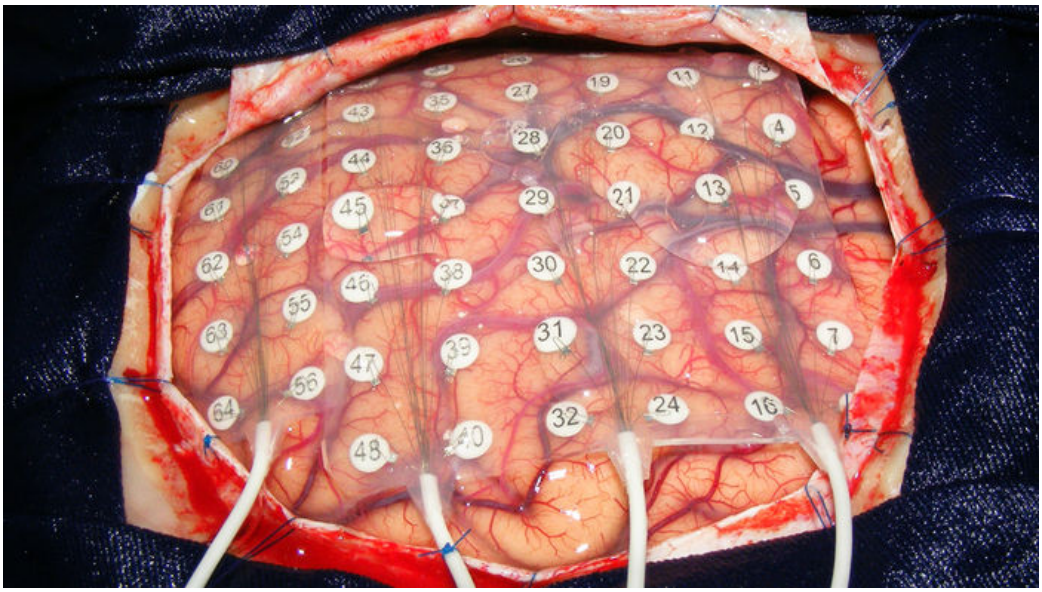

Figure 2.4: Intracranial electroencephalogram (iEEG) acquisition.

from epilepsy caused by focal cortical dysplasia. The dataset consists of the interictal iEEG recorded for two hours with the sampling rate of 2,000 Hz. The label (focal or non-focal) is judged by clinical experts through the visual. Four patients (patient-1,2,3 and 4) with focal (seizure onset zone SOZ) channel number of 3, 3, 7, 9 and randomly select 3, 3, 7, 9 non-focal (non-SOZ) channels, then split the channel data into 20 seconds segments, each patient has 2,160 2,160 5,040 6,480 samples (half of focal and non-focal). An example of focal and non-focal samples are shown in Fig. 2.6. This dataset is recorded under approval from the Juntendo University Hospital Ethics Committee and the Tokyo University of Agriculture and Technology Ethics Committee.

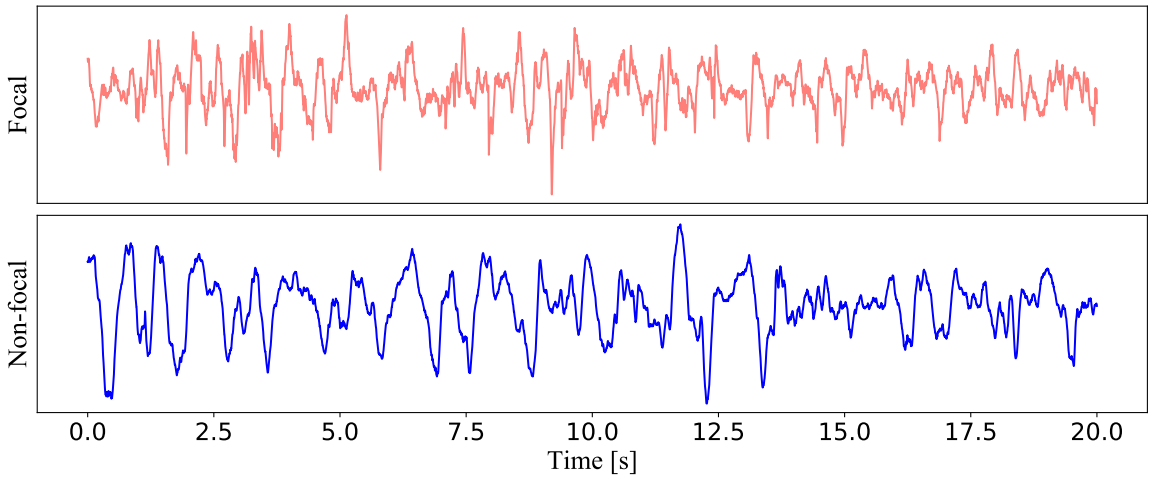

Figure 2.5: Samples of Focal and non-focal iEEG (Bern Barcelona Dataset).

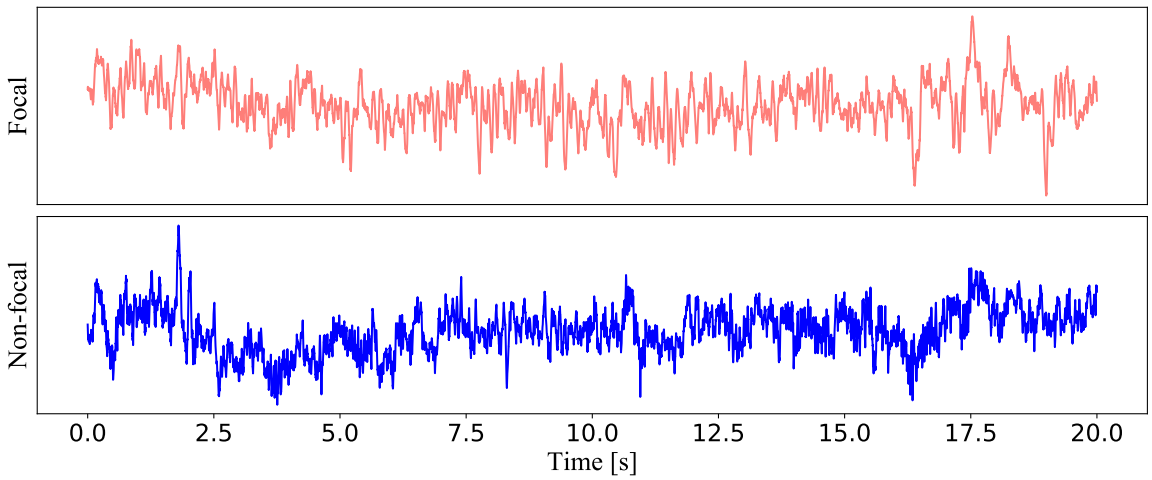

Figure 2.6: Samples of Focal and non-focal iEEG (Juntendo Dataset).

3. Feature Extraction Methods

The brain disorders are usually diagnosed by visual inspection of EEG/iEEG. For epilepsy, EEG are used to diagnose and iEEG are used for epileptic focus localization. EEG/iEEG analysis for assisting in the diagnosis of epilepsy is started in the early [37–42]. Regardless of the EEG and iEEG diagnostic process, clinical experts need long-term visual diagnosis, In this article, a diagnostic system is shown in Fig. 3.1, after record a long time iEEG data, we ask clinical experts to mark a small piece of data and then use this data with label to train a model and predict the remaining data.

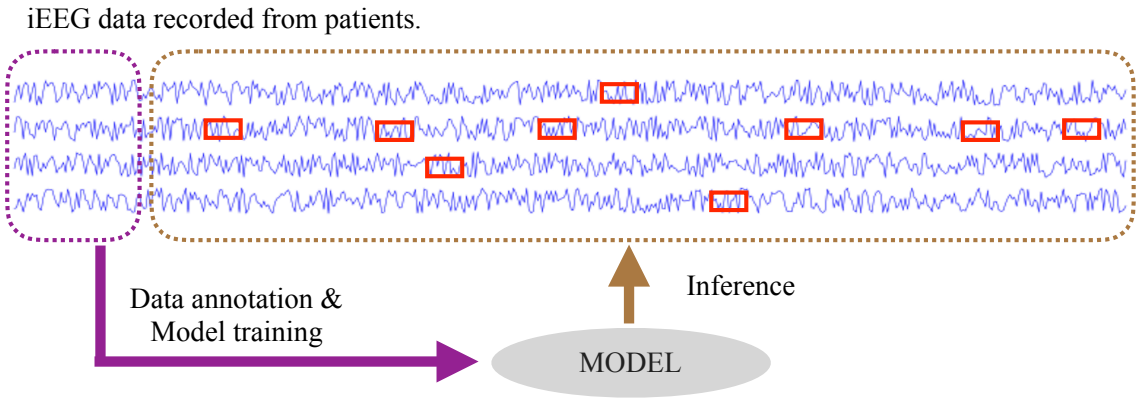

Figure 3.1: Diagnostic system (Only a small piece of iEEG data (purple dotted box) need to be labeled, and the prediction (brown dotted box) include the channel and period information. The red box means that this part of the iEEG data is focal data).

3.1 Feature Extraction by Filter & Entropy

Considering that epilepsy is caused by abnormal discharge of brain cells, we consider using entropy [33] as the method of feature extraction, the flowchart of feature extract procedure is shown in Fig. 3.2.

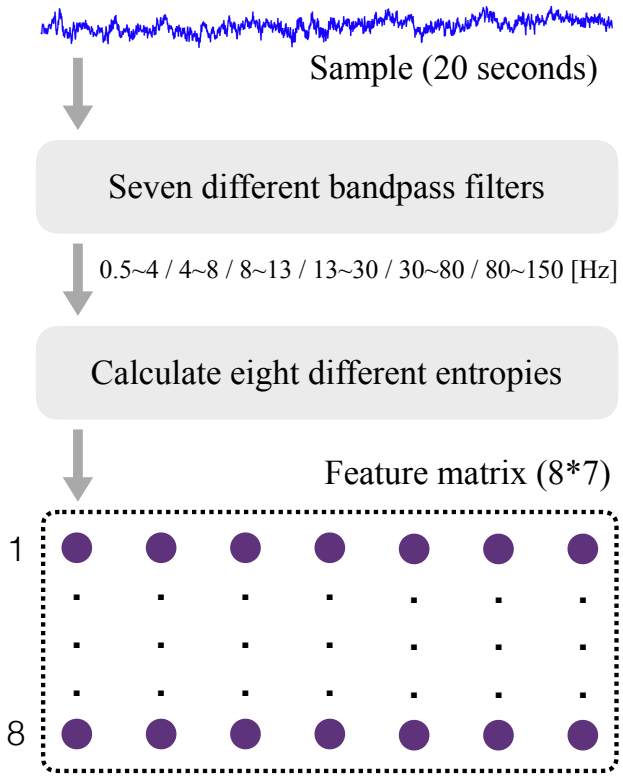

Figure 3.2: Flowchart of feature extract procedure: Entropy & Filter.

3.1.1 Filter

In this method, we extract the features from raw iEEG by using seven different bandpass filters and eight different entropies, corresponding to different brain states, the bandpass filters are shown in Table 3.1, which are the commonly used physiological frequency bands. We use high frequency components (HFO) which shows important roles in iEEG [43]. But in the Bern-Barcelona dataset, because the data is processed by the bandpass filter between 0.5 and 150 Hz, so the bandpass filters used in Bern-Barcelona dataset are Delta (0.5-4 Hz),

Theta (4-8 Hz), Alpha (8-13 Hz), Beta (13-30 Hz), Gamma (30-80 Hz) and Ripple (80-150 Hz). The iEEG processed by filter is shown in Fig. 3.3 and Fig. 3.4

Table 3.1: Bandpass filters based on physiological frequency.

Brain signals	Frequency [Hz]
Delta	0.5 - 4
Theta	4 - 8
Alpha	8 - 13
Beta	13 - 30
Gamma	30 - 80
Ripple	80 - 250
Fast Ripple	250 - 600

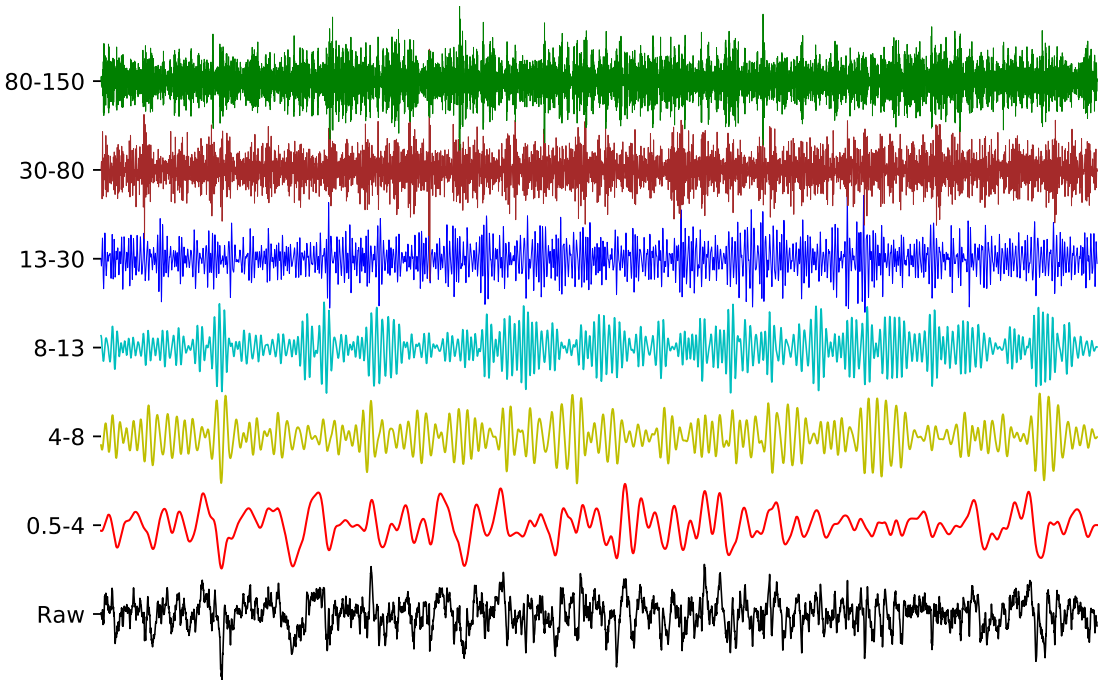

Figure 3.3: Filtered focal signal (Bern Barcelona Dataset).

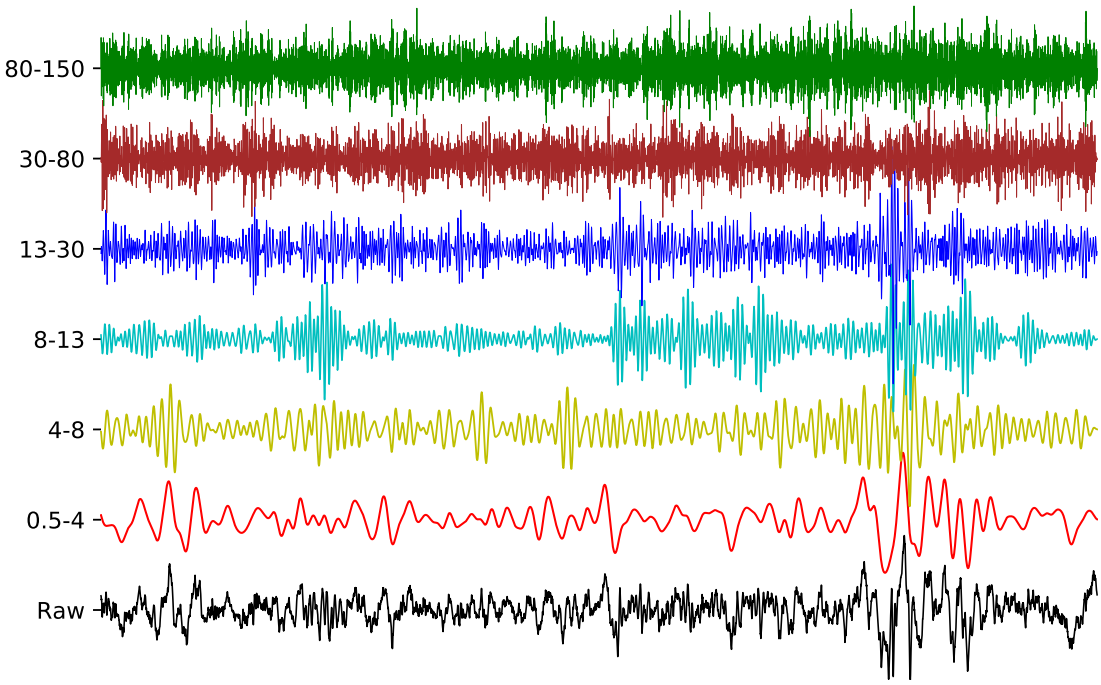

Figure 3.4: Filtered non-focal signal (Bern Barcelona Dataset).

3.1.2 Entropy

Eight different entropies are calculated for each iEEG sample following the bandpass filtering. The entropies are as follows, Shannon entropy, Renee entropy, Generalized entropy [44] [45], Phase entropy (two types) [46], Approximate entropy [47], Sample entropy [48] and Permutation entropy [49]. After the calculation of seven (six for Bern-Barcelona dataset) bandpass filters and eight entropies, we obtain a feature matrix with the size of 8×7 (8×6 for Bern-Barcelona dataset).

3.2 Feature Extraction by Short Time Fourier Transform

Time frequency analysis is a method often used in signal processing, in this way we can analysis signal from another perspective. In clinical practice, epileptic spikes are used for

diagnosis of epilepsy, which can classify and localize the epileptic focus. But in the time domain, because of various noises in the iEEG, spikes are hard to identify, so we try to convert it to the frequency domain, making the focal data feature easier to identify. Ordinary time frequency analysis can not be used for the non-stationary signal because it does not have time resolution. In the article, we use STFT as another method for feature extraction. STFT divide the signal into many short segments, and then compute the Fourier transform for each segment. For a determined signal $x(t)$, the time-frequency domain at each time point can be obtained by the following *(3.1)*.

$$STFT\{x(t)\}(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-j\omega t}\,dt \tag{3.1}$$

where $w(t)$ is the *Hann* window function centered around zero. By using STFT, we can obtain the frequency information in a short period of time. The STFT samples of Bern-Barcelona and Juntendo dataset are shown in Fig. 3.5

3.3 Classifier

3.3.1 Classifier of Support Vector Machine

In the classification step, we are chosen several typical supervised learning methods. Support vector machine (SVM) method [50] mapped the input data to a very high dimensional feature space, then construct a decision surface in the feature space. In this article the polynomial kernel is used.

3.3.2 Classifier of Fully Connected Neural Network

The second classifier is fully connected neural network (FCNN) and the parameters are as follows: the size of input layer is 48 / 56 (Bern-Barcelona / Juntendo), the size of hidden layer one is 32, the size of hidden layer two is 32, the size of output layer is 2, the loss

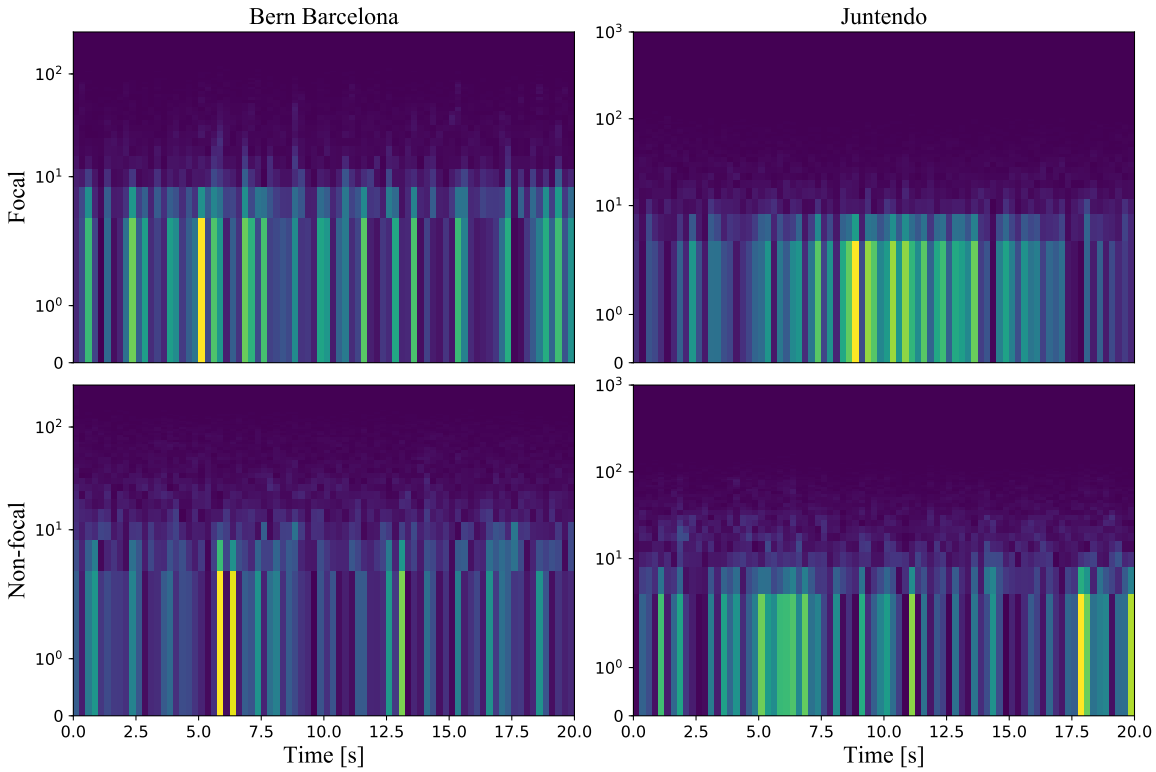

Figure 3.5: Examples of STFT results, left column is Bern-Barcelona data and right column is Juntendo data. window length is one seconds with 80% overlap. For every subfigure, the X axis is 0-20 seconds. the Y axis is 0-256 Hz for Bern-Barcelona data and 0-1,000 Hz for Juntendo data, the Y axis is scale by symlog.

function is binary cross entropy, the optimizer is Adam and the batch size is 64.

3.3.3 Classifier of Convolutional Neural Network

The third classifier is convolutional neural network (CNN) and the parameters are as follows: the size of input image is 8×6 / 8×7 (Bern-Barcelona / Juntendo), the size of convolution kernel is 3×3, the number of convolutional filters is 32, the size of max pooling is 2×2, the size of hidden layer is 128, the size of output layer is 2, the loss function is categorical cross entropy, the optimizer is Adam and the batch size is 64.

3.4 Experimental Result

In this section, we use Bern-Barcelona and Juntendo dataset to evaluate our method, respectively.

3.4.1 Experimental Result of Bern Barcelona Dataset

In Bern-Barcelona dataset, there are 7,500 focal signals and 7,500 non-focal signals, every sample is 20 seconds with the sampling rate of 512 Hz. Because the dataset is patient, channel and time mixed, we use 10-fold cross validation. The results of classification accuracy by FCNN & Entropy are shown in Fig. 3.6, CNN & Entropy are shown in Fig. 3.7 CNN & STFT are shown in Fig. 3.8 and results are shown in Table 3.2.

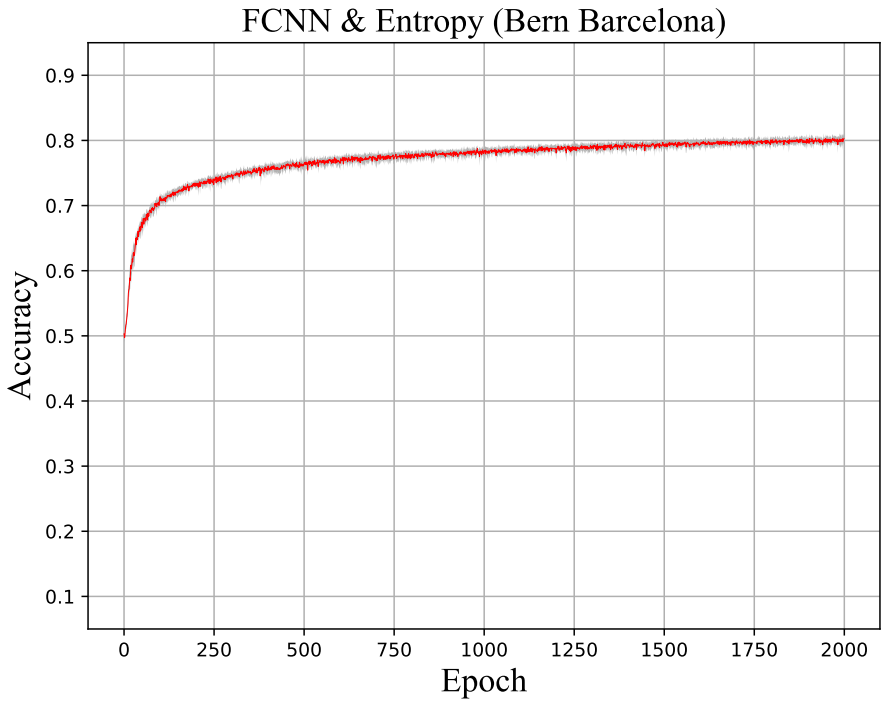

Figure 3.6: Result of FCNN & Entropy (Bern Barcelona Dataset), test accuracy vs. number of epochs. Red line: Average of classification test accuracy (10-folds). Gray area: Standard deviation.

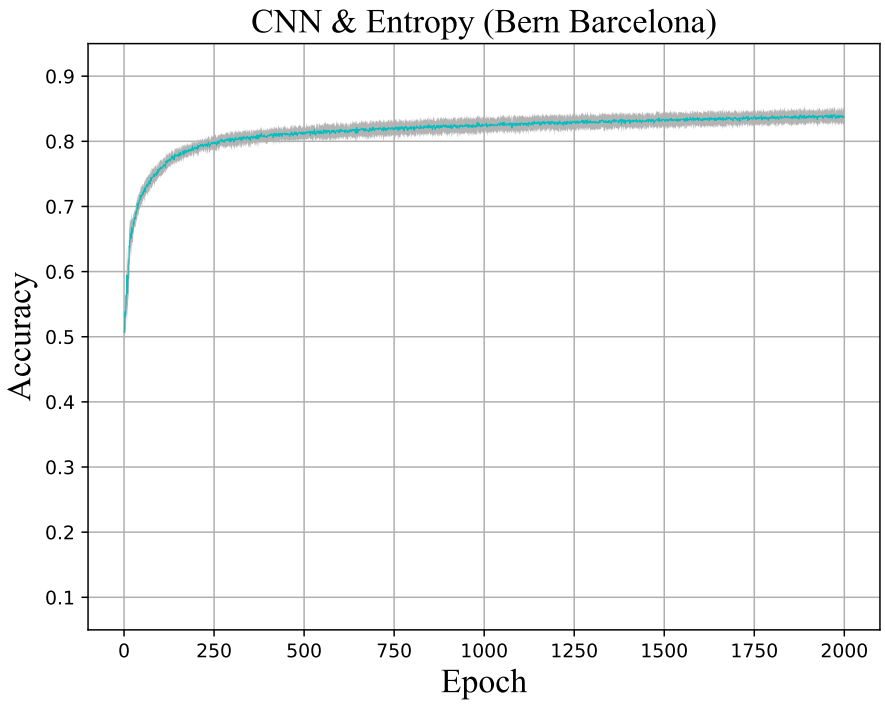

Figure 3.7: Result of CNN & Entropy (Bern Barcelona Dataset), test accuracy vs. number of epochs. Cyan-blue line: Average of classification test accuracy (10-folds). Gray area: Standard deviation.

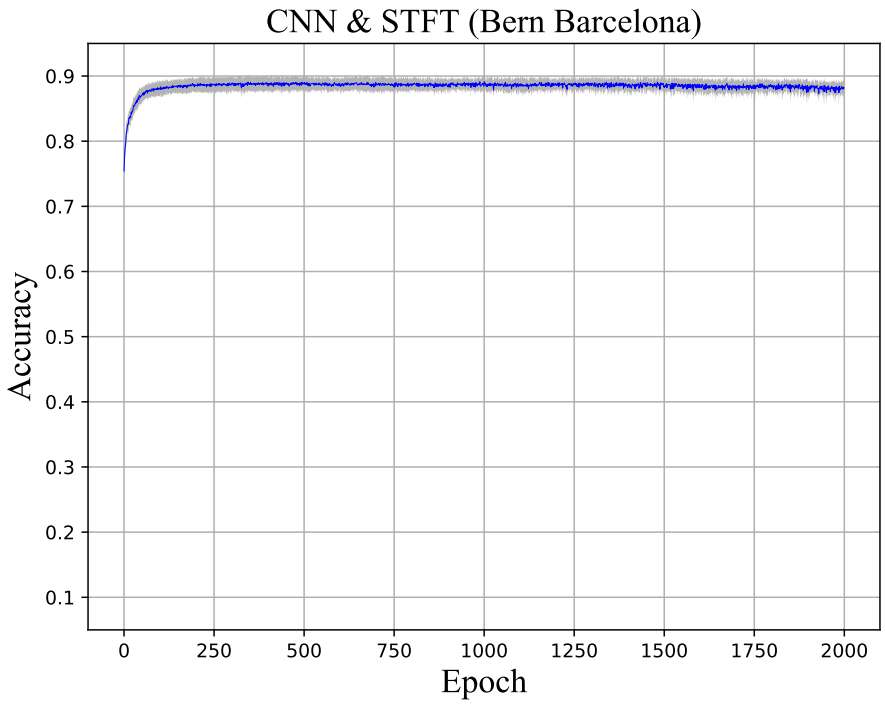

Figure 3.8: Result of CNN & STFT (Bern Barcelona Dataset), test accuracy vs. number of epochs. Blue line: Average of classification test accuracy (10-folds). Gray area: Standard deviation.

Table 3.2: Result of SVM & Entropy, FCNN & Entropy, CNN & Entropy and CNN & STFT (Bern Barcelona Dataset), accuracy [%] over last 10 epochs (Mean $\pm$ Standard deviation).

Accuracy [%]	SVM, Entropy	FCNN, Entropy	CNN, Entropy	CNN, STFT
Bern Barcelona	80.81 $\pm$ 1.60	80.06 $\pm$ 1.58	83.80 $\pm$ 0.11	**88.14** $\pm$ 0.12

3.4.2 Experimental Result of Juntendo Dataset

In Juntendo dataset, there are four patients (Patient-1,2,3 and 4), each patient has 2,160 2,160 5,040 6,480 samples (half of focal and non-focal). Currently, for each patient, we recorded iEEG data for two hours. In order to reduce the workload of clinical experts, we use the former part of the iEEG data as train data (one hour and forty-five minutes) and last fifteen minutes iEEG data as test data. The results of classification accuracy by FCNN &

Entropy are shown in Fig. 3.9, CNN & Entropy are shown in Fig. 3.10 CNN & STFT are shown in Fig. 3.11 and results are shown in Table 3.3.

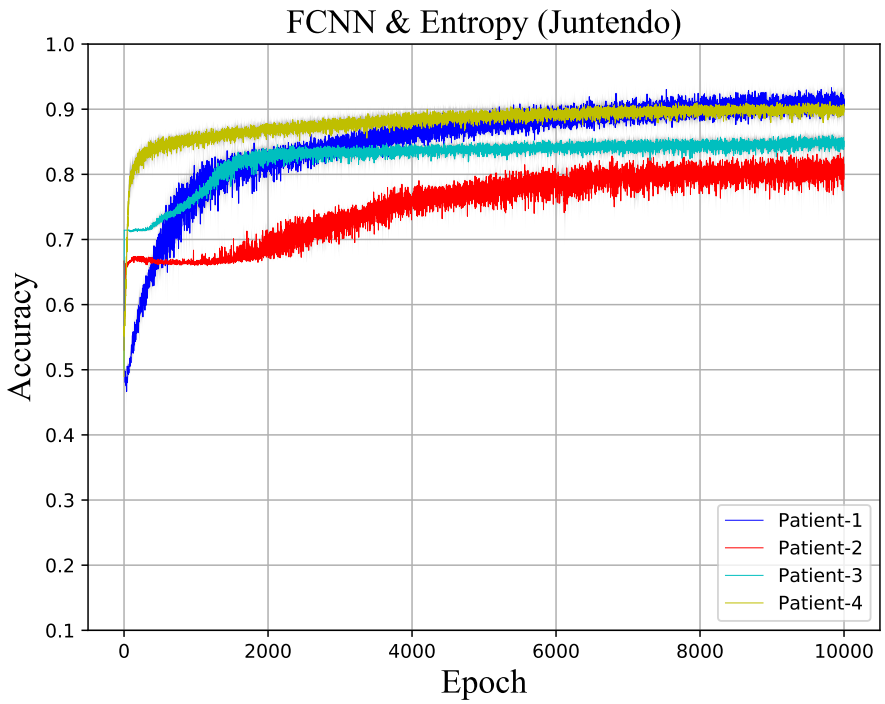

Figure 3.9: Result of FCNN & Entropy (Juntendo Dataset), test accuracy vs. number of epochs. Four different color lines: Average of classification test accuracy (Five repeated experiments). Gray area: Standard deviation.

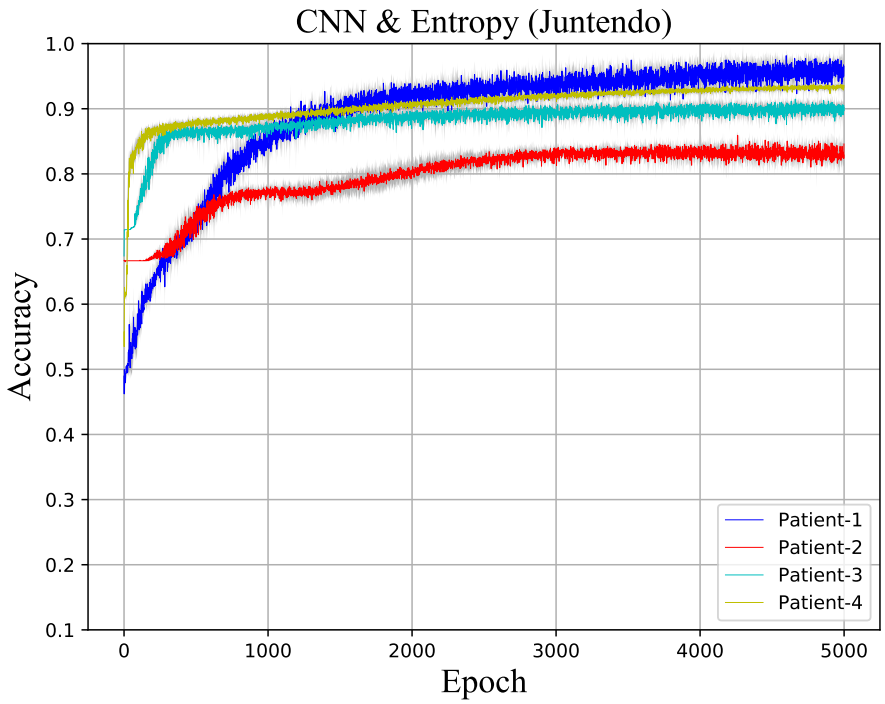

Figure 3.10: Result of CNN & Entropy (Juntendo Dataset), test accuracy vs. number of epochs. Four different color lines: Average of classification test accuracy (Five repeated experiments). Gray area: Standard deviation.

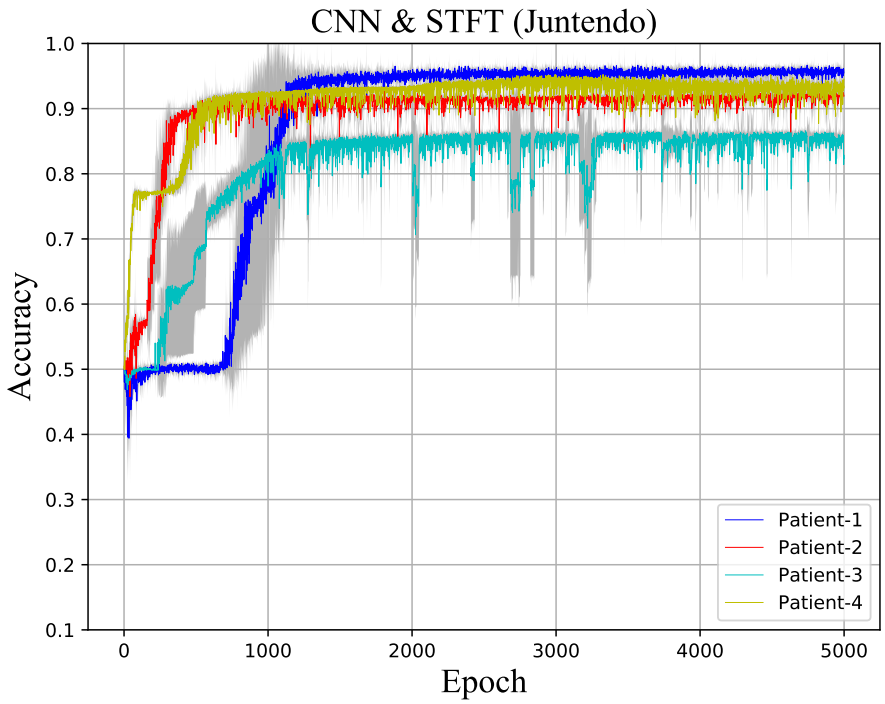

Figure 3.11: Result of CNN & STFT (Juntendo Dataset), test accuracy vs. number of epochs. Four different color lines: Average of classification test accuracy (Five repeated experiments). Gray area: Standard deviation.

Table 3.3: Result of SVM & Entropy, FCNN & Entropy, CNN & Entropy and CNN & STFT (Juntendo Dataset), accuracy [%] over last 10 epochs (Mean $\pm$ Standard deviation).

Accuracy [%]	SVM, Entropy	FCNN, Entropy	CNN, Entropy	CNN, STFT
Patient 1	85.93 $\pm$ 0.00	90.99 $\pm$ 0.68	95.70 $\pm$ 0.65	**95.73** $\pm$ 0.34
Patient 2	84.01 $\pm$ 0.00	81.01 $\pm$ 1.16	83.11 $\pm$ 0.62	**92.11** $\pm$ 0.20
Patient 3	88.57 $\pm$ 0.00	84.74 $\pm$ 0.53	**89.94** $\pm$ 0.51	84.31 $\pm$ 1.48
Patient 4	89.14 $\pm$ 0.00	90.13 $\pm$ 0.34	**93.28** $\pm$ 0.24	93.04 $\pm$ 1.25

4. Deep Learning Methods

In the CNN model, the early layers following input layers are convolutional layers. In convolutional layers, the convolution operation is applied to extract feature maps from the input file of the previous layer, One-dimension convolutional layer: It consists of one-dimension learnable filters which slide across one-dimension input file like time series. Two-dimension convolutional layer: A two-dimension filter is convolved across the width and height of the input file like images. The activation map is obtained by computing the dot product of the input file and the filter. Then after additive bias and non-linear map by activation functions, feature maps of the convolutional layer are outputted to passed to the next layer in the CNN model. Pooling Layer: In the pooling layer, feature maps from the upper layer are down-sampled to reduce the size, lower the calculation complexity and prevent overfitting. In CNN, the pooling layer is a common down-sampling method that the feature maps are separated into many rectangle regions, and then each region features are obtained. Pooling operation various, for instance, max-pooling operation selects only the maximum value in each region, while mean-pooling obtains the mean value of each region. Pooling is the expression of local features and consequently reduces the dimension. Batch Normalization Layer: Batch normalization layer is applied to normalizes the output of the previous layer by subtracting batch mean and dividing by batch standard deviation, to fight the internal covariate shift problem and increase the stability of a neural network.

For the input x obtain from previous layer, the batch normalization layer first calculates the mean $\mu_{\mathscr{B}}$ and variance $\sigma^2_{\mathscr{B}}$ of a mini-batch $\mathscr{B}$ of size m by *(4.1)* and *(4.2)*. Then

normalized values $\overline{x_i}$ are calculated as *(4.3)* where ε is a constant added to the mini-batch variance for numerical stability. Finally, the $\overline{x_i}$ are shifted and scaled as *(4.4)* that the parameters γ and β are to be learned [51] .

$$\mu_{\mathscr{B}} = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{4.1}$$

$$\sigma^2_{\mathscr{B}} = \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_{\mathscr{B}})^2 \tag{4.2}$$

$$\overline{x_i} = \frac{x_i - \mu_{\mathscr{B}}}{\sqrt{\sigma^2_{\mathscr{B}} + \varepsilon}} \tag{4.3}$$

$$y_i = \gamma\overline{x_i} + \beta \tag{4.4}$$

4.1 One-dimensional Convolutional Neural Network

The developed one-dimension convolutional neural network architecture has 27 layers consist of the input, one-dimension convolution, max-pooling, dropout, batch normalization, and fully connected layers. The architecture of 1D-CNN is shown as Fig. 4.1. The convolution layers following the input layer, perform the convolution operation on the input raw iEEG signals. The size of the filter and stride are set as three and two, respectively. Feature maps obtained from the previous convolution process are then successively processed by the max-pooling layer and batch normalization layer. Before fed into the fully connected layers, feature maps are flattened to transform dimension. The sigmoid layer is used in the last layer of the architecture to execute the classification process. In this layer, the input EEG signals are classified as focal or non-focal.

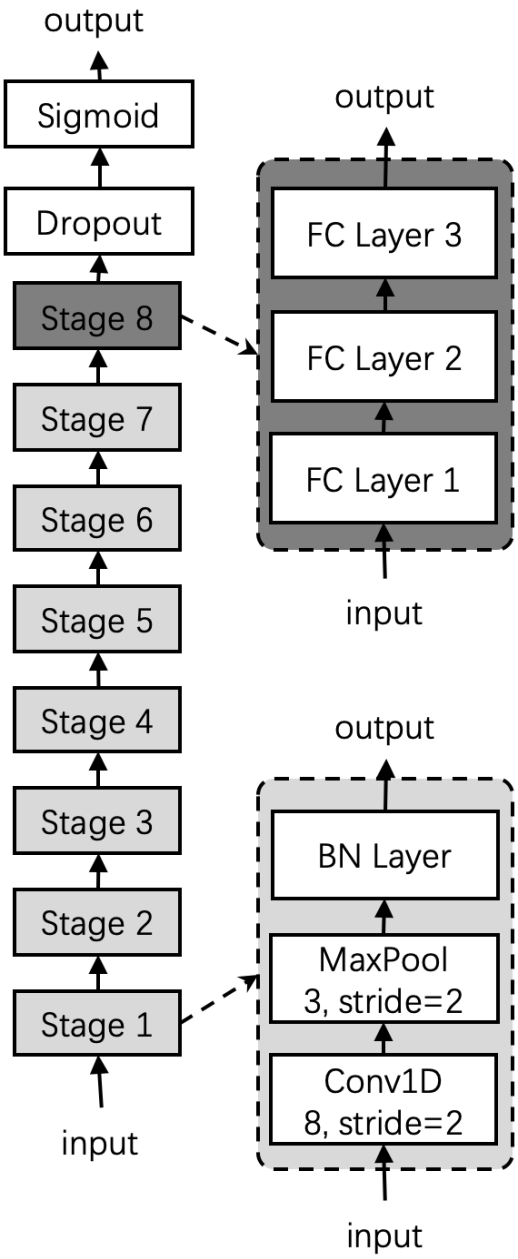

Figure 4.1: Model architecture of 1D-CNN.

4.2 Mixed Convolutional Neural Network

In the previous TFCNN architecture, before feeding into the neural network the signals need to perform extraction and selection of features manually. The most used time-frequency analysis method like STFT has the capability to extract local information at a one-time

scale determined by a single filter, limiting the flexibility of the model. To address this problem, we propose a mixed 1D-2D convolutions model, instead of STFT, we select to setting one-dimension convolution in the earlier layers, because it is easier to optimize the parameter configuration when each layer is treated independently, and it also enables using different input feature maps or receptive field sizes. The architecture of MCNN is shown as Fig. 4.2. The feature maps from one-dimension convolution layers are reshaped and then successively fed to two-dimension convolution layers and fully connected layer to perform further feature extraction and classification.

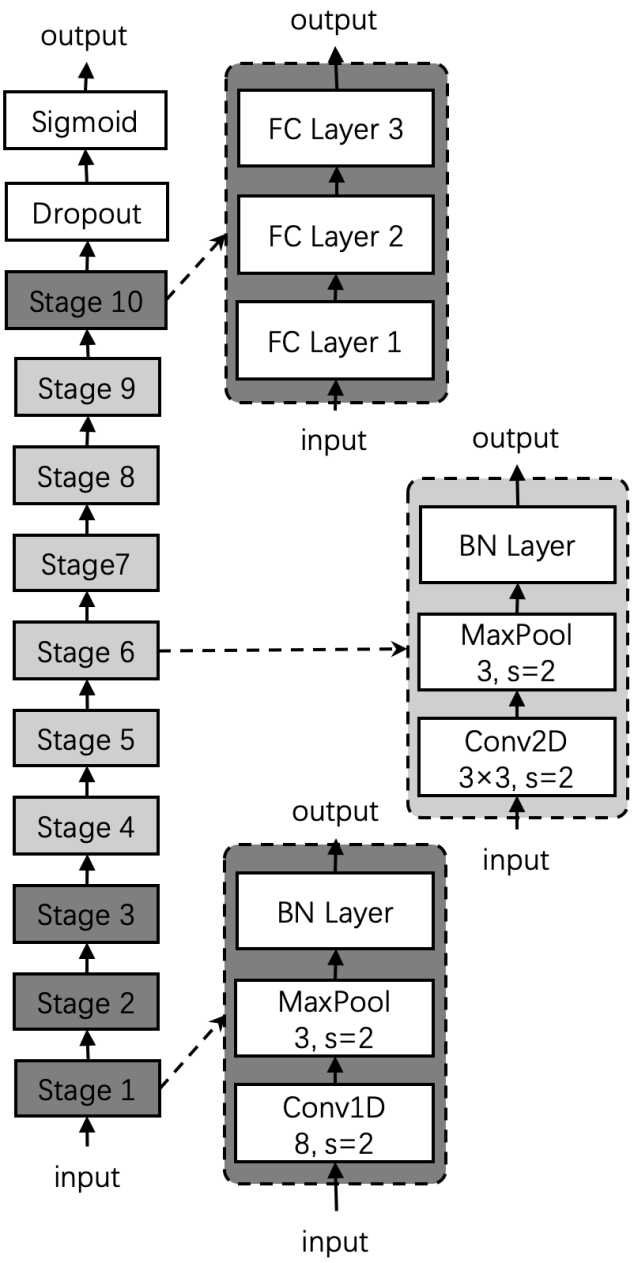

Figure 4.2: Model architecture of MCNN.

4.3 Experimental Result

4.3.1 Experimental Result of One-dimensional Convolutional Neural Network

Bern Barcelona iEEG Dataset has been used to evaluate our proposed models, In Bern-Barcelona dataset, there are 7,500 focal signals and 7,500 non-focal signals, every sample is 20 seconds with the sampling rate of 512 Hz. Because the dataset is patient, channel and time mixed, 10-folds cross-validation is used. In the training stage, the network has been trained to recognize two classes of iEEG signals of focal and non-focal signals. It requires a large number of computational overhead to use one iteration of full training set to perform each epoch, hence stochastic gradient descent (SGD) training is used in this paper. In each epoch of the training, the 13,500 data are randomly divided into 100 batches, which are fed into the network in turn. Training performance was monitored during the training stage until getting the best accuracy on the training set with minimum train loss. The accuracy of the model is shown in Fig. 4.3 and Table 4.1.

4.3.2 Experimental Result of Mixed Convolutional Neural Network

For MCNN model, we also use 10-folds cross-validation method. The accuracy of the model is shown in Fig. 4.4 and Table 4.1. In order to analyze the model performance more comprehensively, we analyzed the confusion matrices for MCNN model, the result is shown in Table 4.2.

Table 4.1: Result of 1D-CNN & MCNN (Bern Barcelona Dataset), accuracy [%] (10-folds) over last 10 epochs (Mean $\pm$Standarddeviation).

Model	Accuracy [%]
1D-CNN	86.39 $\pm$0.069
MCNN	92.80 $\pm$0.083

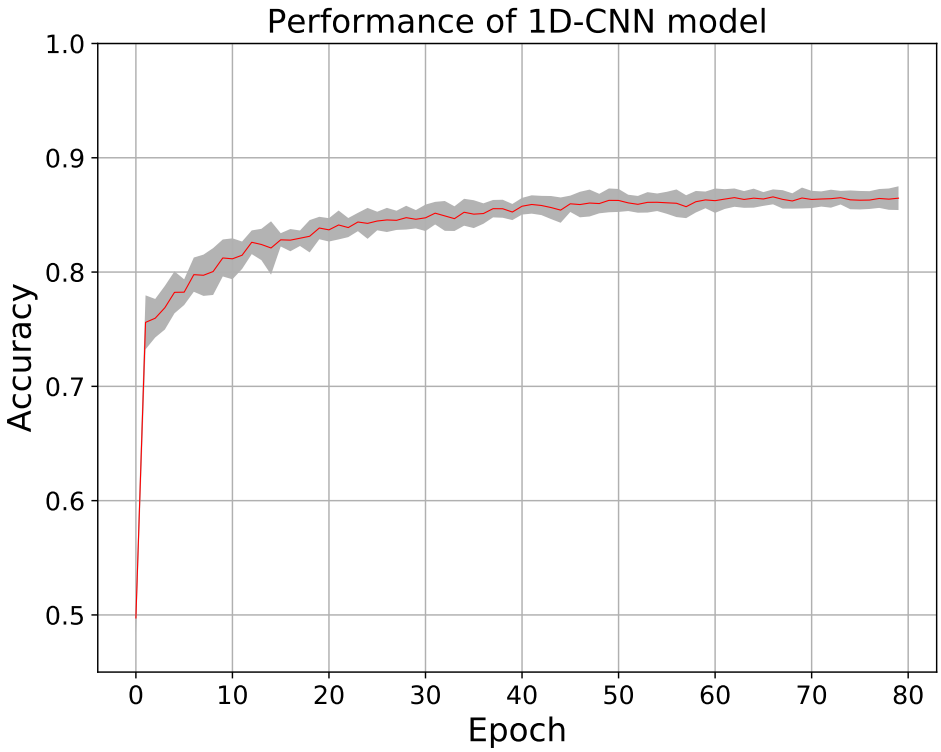

Figure 4.3: Result of 1D-CNN model (Bern Barcelona Dataset), test accuracy vs. number of epochs. Red line: Average of classification test accuracy (10-folds). Gray area: Standard deviation.

Table 4.2: Confusion matrices of classification accuracy for MCNN (Bern Barcelona Dataset).

	Focal (Predict)	Non-focal (Predict)
Focal (True)	687	52
Non-focal (True)	56	705
Precision (%)	Recall (%)	Accuracy (%)
92.5	93.0	92.8

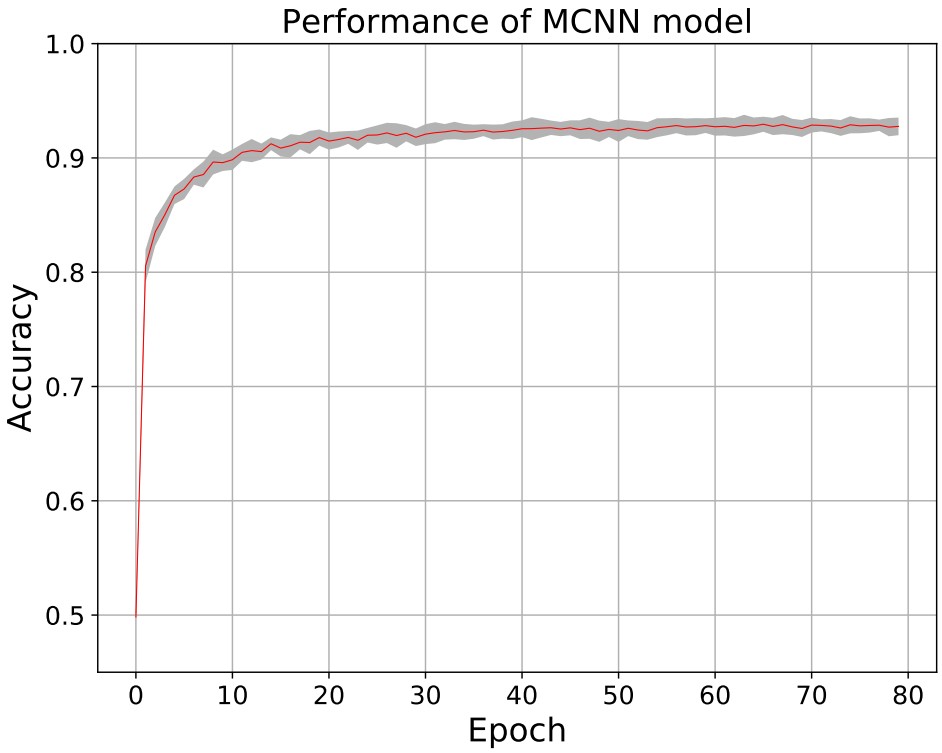

Figure 4.4: Result of MCNN model (Bern Barcelona Dataset), test accuracy vs. number of epochs. Red line: Average of classification test accuracy (10-folds). Gray area: Standard deviation.

5. Weakly Supervised Learning Methods

Up to now, by using supervised learning, we achieved high performance in epileptic focus localization tasks. However, the result is dependent on a large amount of data and labels. It may be available in other fields, but in the medical field, the acquisition of high-quality data with the label is very difficult and expensive. Aiming at the practical problems, we used the PU learning for epileptic focus localization. In the previous chapters, both traditional supervised learning and end-to-end neural network models. Our input is a lot of data and corresponding tags. This requires us to do a lot of labeling work beforehand. Obviously, in the medical field, this is difficult to achieve. For epilepsy iEEG data, current clinical experts are labeled by visual judgment. This has a need to reduce the workload of clinical expert data annotation.

In the figure of Performance comparison of different learning models 5.1, we compared several common learning models and there is a balance between model performance and data annotation cost. There is a balanced approach between model performance and data tagging: semi-supervised learning. The difference between supervised learning and semi-supervised learning is shown in Fig 5.2. Compared to supervised learning, semi-supervised learning does not require all data to be labeled. It only requires that some of the data in each class of data have the label. This greatly reduces the workload of annotation for clinical experts. In order to further improve the work of clinical experts, we use PU learning. In

contrast to semi-supervised learning, PU learning requires only one class of data with label. This further reduces the amount of annotation work. Below we will detail the model of the PU learning method, mathematical proof and experimental results.

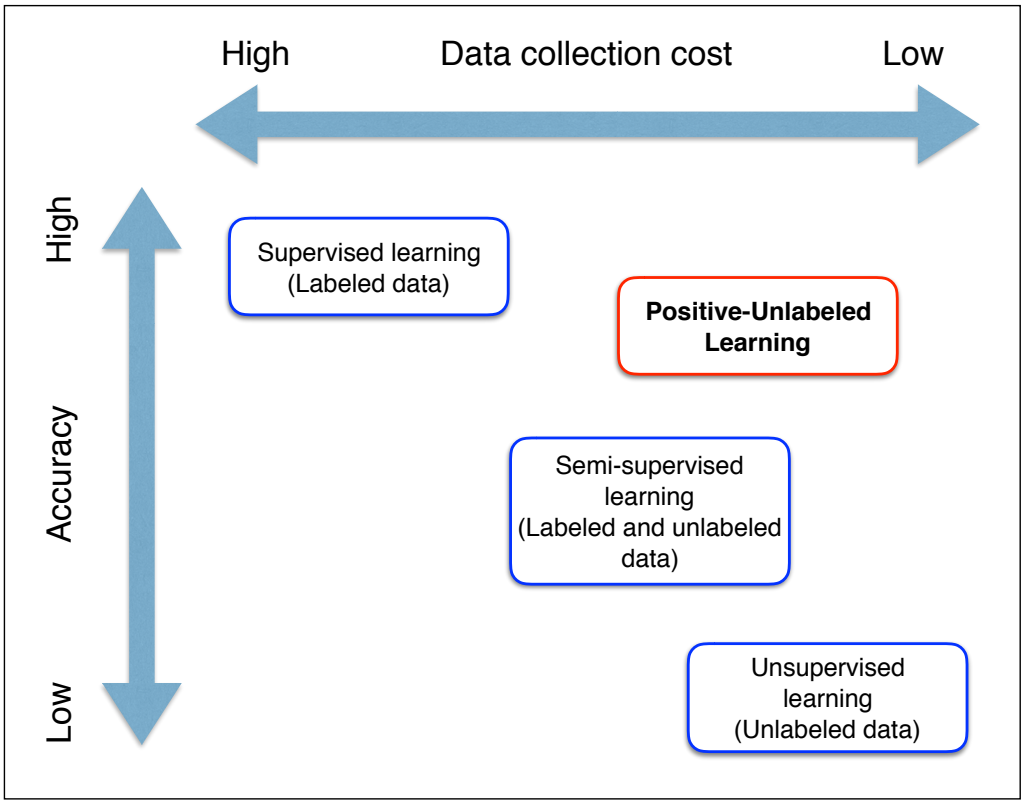

Figure 5.1: Performance and data annotation cost comparison of different learning models (Include Supervised learning, Positive Unlabeled learning, Semi-supervised learning and Unsupervised learning).

5.1 Positive Unlabel Learning

There is a definition of PU learning [52]: Given a set of examples of an particular class P (called the positive class) and a set of unlabeled examples U, which contains both class P and non-class P (called the negative class) instances, the goal is to build a binary classifier to classify the test set T into two classes, positive and negative, where T can be U.

In the Fig. 5.3, the left part is the traditional model of supervised learning, we have two types of data and every data has label, in the right part, we have two kind of data, one is a

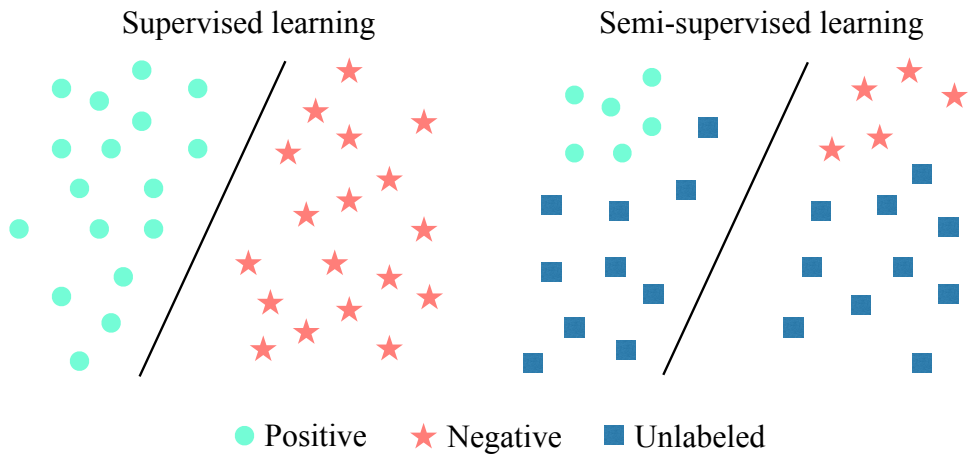

Figure 5.2: Model comparison of supervised learning and semi-supervised learning.

small amount of positive date with label, the other is data without any label.

There are some similarities between the PU learning and semi-supervised learning, which is also used in medical field [53–55], but unlike to the semi-supervised learning, PU learning only need give label to one kind of data.

According to how to solve the problem of a kind of data without label, the PU learning can be divided into two categories. One category is try to find the reliable negative data (RN) data from unlabeled data, then by using the postivite data and RN data, the PU learning can be regarded as a binary classifier [56] [52] [57]. The other category is regarding unlabeled data as negative data, and the weight of negative data is adjusted in the loss function [58–60], our article uses this method. Detailed description and mathematical are as follows:

Problem settings: Let $\mathbf{x}$ be the input vector which is calculate by filter and entropy or STFT from a 20 seconds iEEG data, and $y \in \{\pm 1\}$ be the class label, in the article, $+1$ delegate focal data and -1 delegate non-focal data. The class conditional distributions of focal data and non-focal data are denoted by $p_p(\mathbf{x})$ and $p_n(\mathbf{x})$ respectively, and $p_p(\mathbf{x})$ and $p_n(\mathbf{x})$ are defined by

$$\begin{aligned} p_p(\mathbf{x}) &= p(\mathbf{x} \mid y = +1), \\ p_n(\mathbf{x}) &= p(\mathbf{x} \mid y = -1). \end{aligned} \tag{5.1}$$

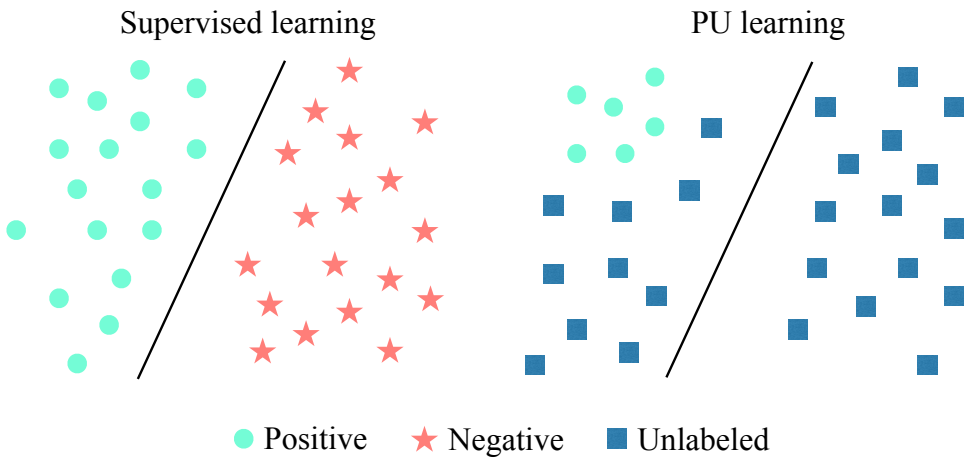

Figure 5.3: Model comparison of supervised learning and PU learning.

The prior probabilities for focal data and non-focal data are denoted by $\pi_p = p(y = +1)$ and $\pi_n = p(y = -1)$ and $\pi_n = 1 - \pi_p$, in this article, the π_p is assumed known in advance. The marginal distribution of unlabeled data (i.e. the distribution of focal and non-focal data) is defined by

$$p(\mathbf{x}) = \pi_p p_p(\mathbf{x}) + \pi_n p_n(\mathbf{x}). \tag{5.2}$$

Risk estimators: we use the empirical unbiased risk estimator that is proposed by [61] [62] [63], the $g(\cdot)$ denotes the binary classification function and $\ell(g(\mathbf{x}), \pm 1)$ is the loss function. Therefore, the $\hat{R}_p^+(g)$ and $\hat{R}_n^-(g)$ denotes the empirical risks for focal and non-focal data, respectively. The empirical risk for focal data, i.e., $\hat{R}_p^+(g)$ is calculated by (5.3) and the empirical risk for non-focal data, i.e., $\hat{R}_n^-(g)$ is calculated by (5.4), then the risk estimator is defined by (5.5).

$$\hat{R}_p^+(g) = \mathbb{E}_{\mathbf{x} \backsim p_p(\mathbf{x})} \ell(g(\mathbf{x}), +1), \tag{5.3}$$

$$\hat{R}_n^-(g) = \mathbb{E}_{\mathbf{x} \backsim p_n(\mathbf{x})} \ell(g(\mathbf{x}), -1). \tag{5.4}$$

$$\hat{R}_{pn}(g) = \pi_p \hat{R}_p^+(g) + \pi_n \hat{R}_n^-(g), \tag{5.5}$$

Because we only have the label of focal data, the distribution of non-focal data is unknown. in equation (5.4), $\hat{R}_n^-(g)$ can not be computed straightforwardly. However, by using the equation (5.2), the distribution of non-focal data can be represented by

$$\pi_n p_n(\mathbf{x}) = p(\mathbf{x}) - \pi_p p_p(\mathbf{x}). \tag{5.6}$$

Hence, the empirical risk for non-focal data can be computed by

$$\pi_n \hat{R}_n^-(g) = \hat{R}_u^-(g) - \pi_p \hat{R}_p^-(g), \tag{5.7}$$

where $\hat{R}_u^-(g)$ and $\hat{R}_p^-(g)$ are the empirical risks under the distribution of unlabeled data and focal data, respectively, which are defined by

$$\begin{aligned} \hat{R}_u^-(g) &= \mathbb{E}_{\mathbf{x} \backsim p(\mathbf{x})} \ell(g(\mathbf{x}), -1), \\ \hat{R}_p^-(g) &= \mathbb{E}_{\mathbf{x} \backsim p_p(\mathbf{x})} \ell(g(\mathbf{x}), -1). \end{aligned} \tag{5.8}$$

Finally, the risk estimator in equation (5.5) can be approximated indirectly by

$$\hat{R}_{pu}(g) = \pi_p \hat{R}_p^+(g) + \hat{R}_u^-(g) - \pi_p \hat{R}_p^-(g). \tag{5.9}$$

In general, $g(\mathbf{x})$ can be any classifier functions, such as linear discriminative analysis (LDA), SVM, FCNN and so on. Due to the recent great success of neural networks, in this study, we employ a three layers FCNN as the binary classifier function $g(\mathbf{x})$. Based on the objective function shown in equation (5.9), we can thus easily employ the BP algorithm to learn the deep neural network for PU problem.

5.2 Experimental Result

We use Bern Barcelona dataset and Juntendo dataset to evaluate PU learning, respectively.

5.2.1 Experimental Result of Bern Barcelona Dataset

we use Bern-Barcelona dataset to evaluate our method. In Bern-Barcelona dataset, there are 7,500 focal signals and 7,500 non-focal signals, every sample is 20 seconds with the sampling rate of 512 Hz. Because the dataset is patient, channel and time mixed, we use 10-fold cross validation. In PU learning experimental, 2,142 focal samples are selected as labeled data. the results of classification accuracy are shown in Fig. 5.4 and Table 5.1.

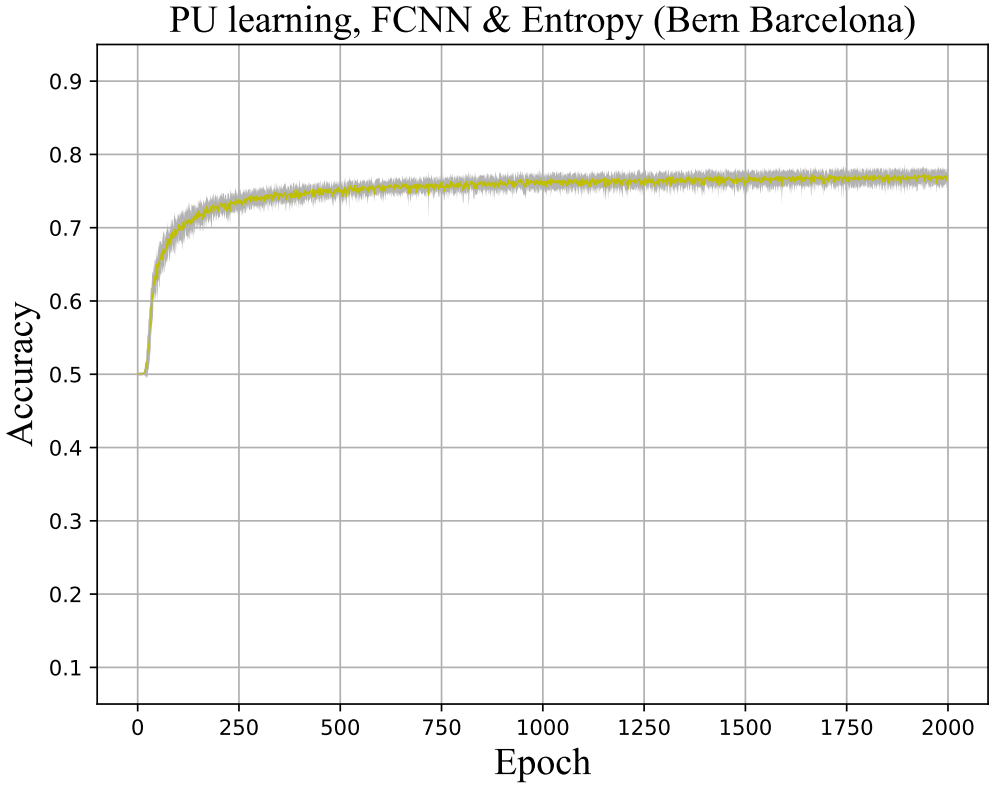

Figure 5.4: Result of PU learning (Bern Barcelona Dataset), test accuracy vs. number of epochs. Yellow line: Average of classification test accuracy (10-folds), Gray area: Standard deviation.

Table 5.1: Result of PU learning (Bern Barcelona Dataset), accuracy [%] over last 10 epochs (Mean $\pm$ Standard deviation).

Accuracy [%]	FCNN & Entropy	PU learning
Bern Barcelona	80.81 ± 1.60	76.91 ± 0.22

5.2.2 Experimental Result of Juntendo Dataset

In Juntendo dataset, there are four patients (Patient-1,2,3 and 4), each patient has 2,160 2,160 5,040 6,480 samples (half of focal and non-focal). The first 315 minutes as train data, and the other as test data. For each patient, we recorded iEEG data for two hours. In order to reduce the workload of clinical experts, we use the former part of the iEEG data as train data (one hour and forty-five minutes) and last fifteen minutes iEEG data as test data. In the PU learning method, we randomly select 300, 300, 700, 900 (patient 1-4) focal sample with label, and all the other data are treated as unlabeled data, the results of classification accuracy are shown in Fig. 5.5 and Table 5.2.

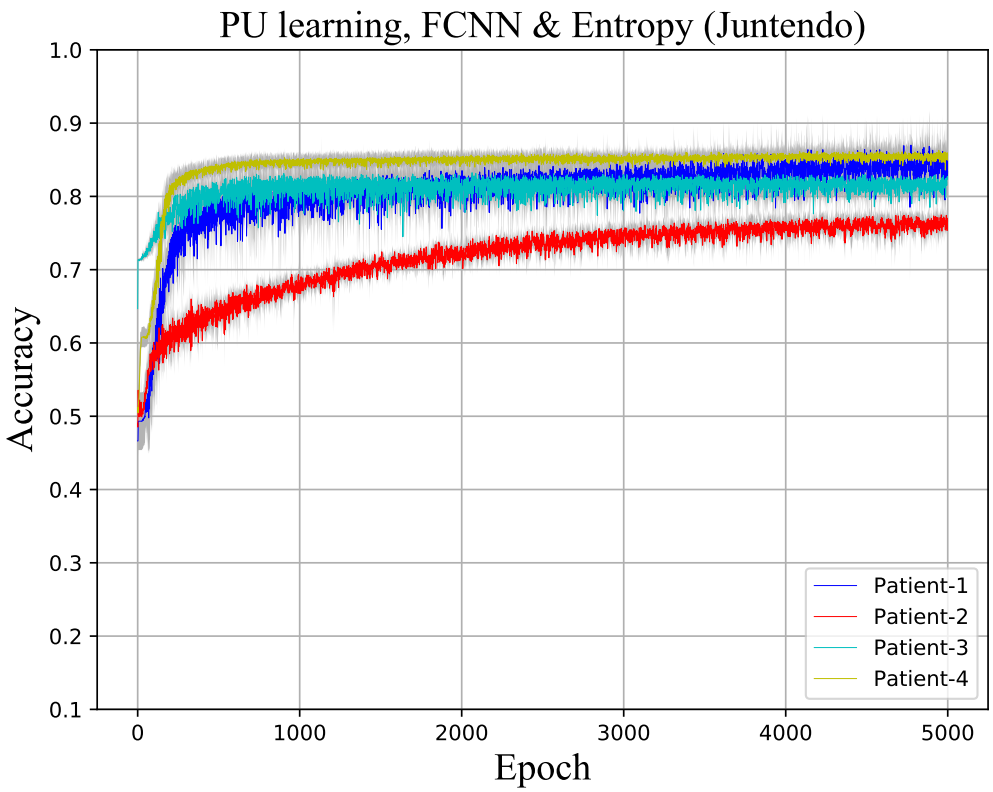

Figure 5.5: Result of PU learning (Juntendo Dataset), test accuracy vs. number of epochs. Four different color lines: Average of classification test accuracy for each patient (Five repeated experiments). Gray area: Standard deviation.

Table 5.2: Result of PU learning (Juntendo Dataset), accuracy [%] over last 10 epochs (Mean $\pm$ Standard deviation).

Accuracy [%]	FCNN & Entropy	PU learning
Patient 1	90.99 ± 0.68	84.09 ± 1.35
Patient 2	81.01 ± 1.16	76.48 ± 0.64
Patient 3	84.74 ± 0.53	82.19 ± 0.43
Patient 4	90.13 ± 0.34	85.47 ± 0.37

6. Data Augmentation Methods

In the aforementioned methods achieve quite good performance, but there is still a problem, the high performance relies on a massive amount of high quality labeled data. In the medical field, both a massive number of data and high-quality labels are often difficult to obtain. To address this problem, we get inspiration from the field of computer vision. In the field of computer vision, data augmentation is a common method of use. Data augmentation is a strategy to increase the amount and diversity of data available for training models without the need to collect new data, therefore we also avoid the data annotation workload. In this paper, we introduce the data augmentation method originally presented in [64] [65] to generate artificial data. The time-domain data is converted to the frequency domain by discrete cosine transform (DCT), and new artificial data is generated by combining different frequency bands from different data and converted back to time-domain data. With the help of the data augmentation method, the result of the model has been improved by about 3%. Our results show that data augmentation is also an effective method to improve performance for time-domain data (iEEG) when there is only limited data.

6.1 Discrete Cosine Transform Based Data Augmentation Methods

To improve the classification performance, we introduce a DCT-based augmentation method, following the same strategy already used in [64, 65]. DCT is often used in signal process-

ing because it has a strong energy compaction property. Similar to the Discrete Fourier Transform (DFT), the DCT is a Fourier-based transform, but using only the cosine function (real part of the complex exponential function). Using the DCT, the time domain signal can be split into a sum of cosine functions ranging from-high frequency to low-frequency. On the other hand, the cosine functions can be reconstructed with the inverse discrete cosine transform. Our model constructs a synthetic sample by exploiting information in the transformed domain which can be specifically described as follows.

Given a random batch of N independent and identically distributed (*i.i.d.*) samples $\{\mathbf{x}_i \in \mathbb{C}^m\}_{i=1}^N$ drawn from an unknown distribution $\mathscr{D}$, we generate a synthetic sample $\mathbf{g}$ by the following model

$$\mathbf{g} = \sum_{i=1}^{N} \Psi_i(\mathbf{x}_i) \in \mathbb{C}^m. \tag{6.1}$$

By restricting $\{\Psi_i(\cdot)\}$ to be linear functions, the synthetic sample $\mathbf{g}$ can be linearly constructed from original samples $\{\mathbf{x}_i\}$. We further assume each Ψ_i is the compound of a sample specific linear transformation $\mathbf{F}_i$ and a sample independent (inverse) transformation $\mathbf{F}^{-1}$ as $\Psi_i = \mathbf{F}^{-1} \circ \mathbf{F}_i, \quad \forall i = 1,2,\cdots,N$. For example, $\mathbf{F}_i$ can model the process of extracting several DFT (or DCT) coefficients from sample $\mathbf{x}_i$, whereas $\mathbf{F}^{-1}$ models the inverse DFT (or inverse DCT) on the extracted coefficients. In our setting, we let the length m of sample $\mathbf{x}_i$ satisfy $m = \sum_{i=1}^N m_i$, where m_i is the width of a segmental coefficients extracted from sample $\mathbf{x}_i$ in the transformed doamin. Without loss of generality, let $\mathbf{F}$ denote a unitary matrix representing some linear transformation (like DFT/DCT) which has orthogonal rows $\mathbf{f}_j \in \mathbb{C}^{1\times m},\ j = 1,2,\cdots,m$:

$$\mathbf{F} = [\mathbf{f}_1;\cdots;\mathbf{f}_{m_{i-1}^+};\mathbf{f}_{m_{i-1}^+ +1};\cdots;\mathbf{f}_{m_i^+};f_{m_i^+ +1};\cdots;\mathbf{f}_m] \in \mathbb{C}^{m\times m},$$

where $m_i^+ = \sum_{k=1}^i m_k$, $i = 1,2,\cdots,N$. Then, $\mathbf{F}_i$ can exact coefficients of the i-th segment by letting

$$\mathbf{F}_i = [\mathbf{0};\cdots;\mathbf{0};\mathbf{f}_{m_{i-1}^+ +1};\cdots;\mathbf{f}_{m_i^+};\mathbf{0};\cdots;\mathbf{0}] \in \mathbb{C}^{m\times m}.$$

Hence, Model (6.1) can be specified as $\mathbf{g} = \sum_{i=1}^{N} \mathbf{F}^{\mathsf{H}} \mathbf{F}_i \mathbf{x}_i$.

Theorem 1 *Let $\mathbb{E}_{\mathbf{x}}$ and $\Sigma_{\mathbf{x}}$ be the expectation and the covariance matrix of $\mathscr{D}$, respectively. Then it holds that*

(I) The expectation of $\mathbf{g}$ satisfies $\mathbb{E}[\mathbf{g}] = \mathbb{E}_{\mathbf{x}}$.

(II) The covariance $\Sigma_{\mathbf{g}}$ of $\mathbf{g}$ satisfies:

$$\Sigma_{\mathbf{x}} - \Sigma_{\mathbf{g}} = \mathbf{F}^{\mathsf{H}} \Big(\sum\nolimits_{i \neq j} \sum\nolimits_{j} \mathbf{F}_i \Sigma_{\mathbf{x}} \mathbf{F}_j^{\mathsf{H}} \Big) \mathbf{F}. \tag{6.2}$$

Theorem 1 indicates the data augmentation process does not change the expectation of the unknown distribution $\mathscr{D}$ and the shift of covariance has an explicit expression specified by the transformation $\mathbf{F}$. Note that since it is impossible to exactly estimate a generally full-rank covariance matrix $\Sigma_{\mathbf{x}}$ from very limited random samples $\{\mathbf{x}_i\}$, the phenomenon of covariance-shift is unavoidable in principle.

The data augmentation proposed workflow Fig. 6.1 is as follows: (1) Randomly choose seven signals (channels) from the dataset (focal and non-focal signals are operated separately) and apply the transform $\mathbf{F}$ (DCT); (2) Using the frequency bands (Delta: 0-4, Theta: 4-8, Alpha: 8-13, Beta: 13-30, Gamma: 30-80, Ripple: 80-150 and Fast Ripple: 150 $\sim$), extract one frequency band of each of the decompositions, from highest to lowest frequencies, and merge the seven extracted parts (frequency bands) to create a new artificial signal. (3) Artificial signals in frequency domain will be transformed back to time domain applying the inverse transform $\mathbf{F}^{-1}$ (IDCT); (4) Finally, the artificial signals are processed using a (three-order Butterworth) bandpass filter between 0.5 and 150 Hz.

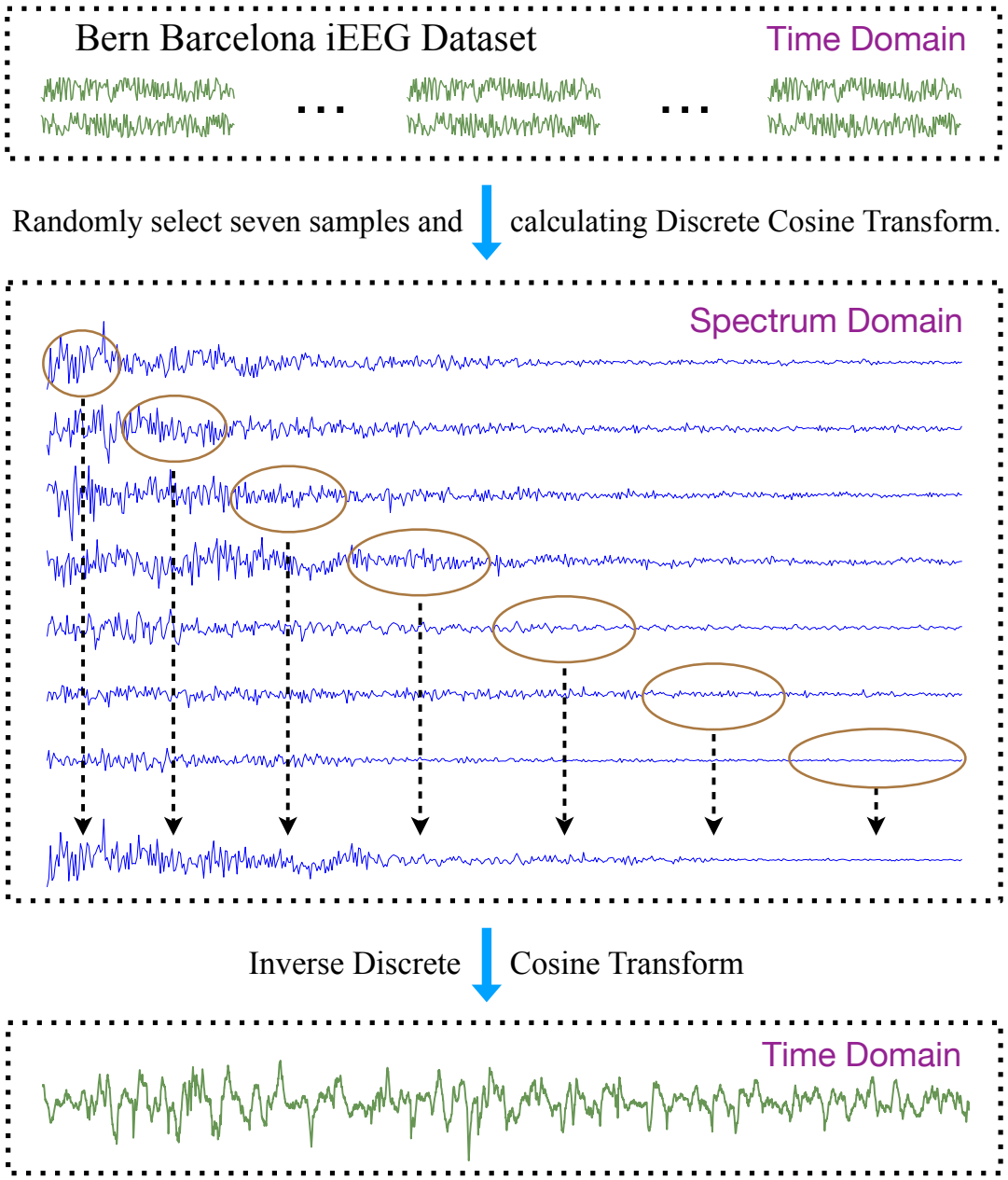

Figure 6.1: Flow chart of the data augmentation method.

6.2 Experimental Result

In our model, we used three types of layers: 1D-CNN layer, pooling layer and fully connected layer. The pooling layer is used to shorten signal length, highlight features and reduce calculation time In addition,it can improve spatial invariance to some extent, such as

translation invariance, scale invariance and deformation invariance. Fully connected layer is used as a decision layer at the end of the network, which can synthesize all the features to generate a conclusion. Our model architecture is as follows: Conv (kernel size = 1×10, number = 32, strides = 1), Conv (1×10, 64, 1), Maxpool (pool size = 1×5, strides = 4), Conv (1×10, 64, 1), Maxpool (1×5, 4), Conv (1×10, 32, 1), Maxpool (1×5, 4), fully connected, a total of eight layers.

First, the original Bern-Barcelona dataset is used to evaluate the 1D-CNN model using a 10-fold cross-validation strategy. These results, shown in Fig. 6.2, will be used with two purposes: (i) to be compared with results already published using the same dataset but different classification models; and (ii) as a benchmark for comparison when using data augmentation technique.

Results shown in Fig. 6.2 have a mean test accuracy of 89.28% with a standard deviation of 0.91, outperforming all works already published using the same datased. Table **??** provides a comparison between the previous work and the proposed method.

In real-world scenarios, it is sometimes difficult to obtain a large amount of medical data with high quality labels, and we often end up with a reduced dataset. To explore this problem, we randomly select a smaller set of test and training data from the Bern-Barcelona dataset, and generated artificial data for the training step, with the aim of investigating whether we can still obtain a good classification model. To do that, we randomly selected 3,000 samples (1,500 focal and 1,500 non-focal signals) and 1,000 samples (500 focal and 500 non-focal signals) as the raw training set and test set, respectively (there is no intersection between the two sets). Using the data augmentation method, we generated 3,000, 6,000 and 9,000 artificial data (artificial focal/non-focal data were generated only by focal/non-focal data) from the raw training set. Then, we trained the 1D-CNN model with raw training set combining different amounts of artificial data. The results are shown in Fig. 6.3 and average test accuracy is shown in Table 6.1.

From these results, we observe that having a small dataset decreases the test accuracy

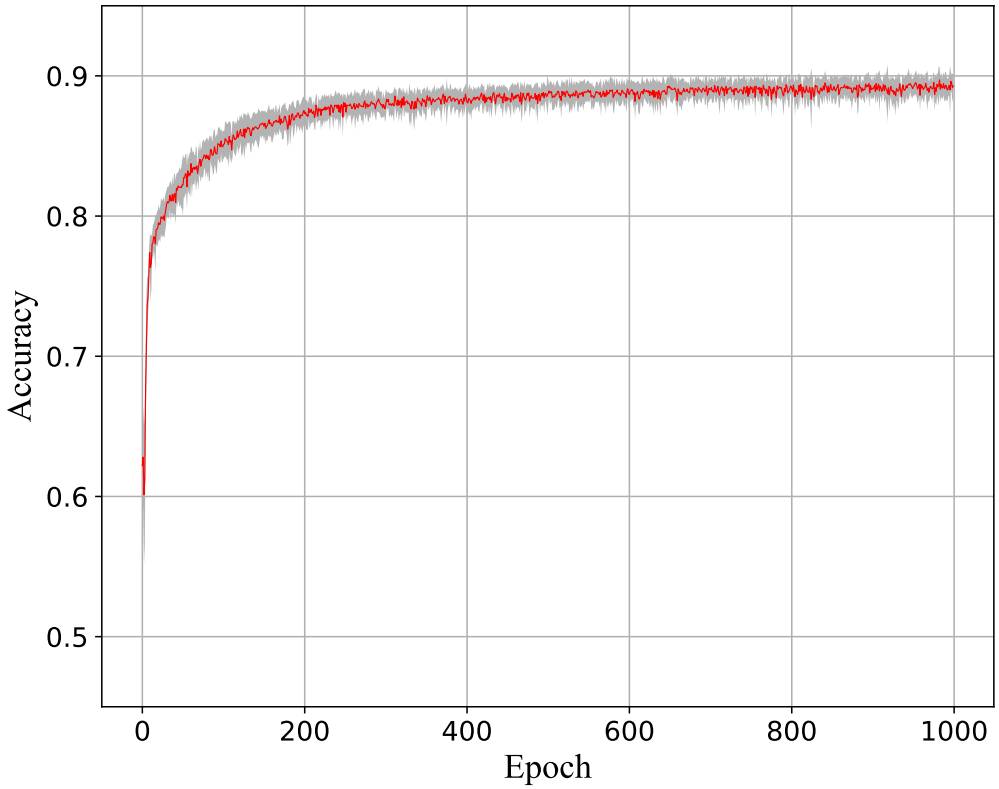

Figure 6.2: Results of the 1D-CNN model with Bern-Barcelona dataset, test accuracy vs. number of epochs. Red line: Average of classification test accuracy (10-folds), Gray area: Standard deviation.

more than 7% using our proposed 1D-CNN model. However, when applying the data augmentation technique, the test accuracy increases gradually up to 83.91% when using 9,000 artificial data together with the original 3,000 raw data. Note that this result is above or similar to most of the ones reported previously but using less than 50% of the real data.

In summary, we explore a deep learning method to avoid the computationally demanding feature extraction step of the classical machine learning methods. Therefore, the 1D-CNN method is applied for the epileptic focus localization prob- lem. The proposed approach can avoid cumbersome feature extraction processes and experimental results show that this approach is effective for this application. In addition, con- sidering the limited amount of medical data in many real-life scenarios, we proposed a data augmentation method. Us- ing artificially generated data, model performance improves without increasing the workload

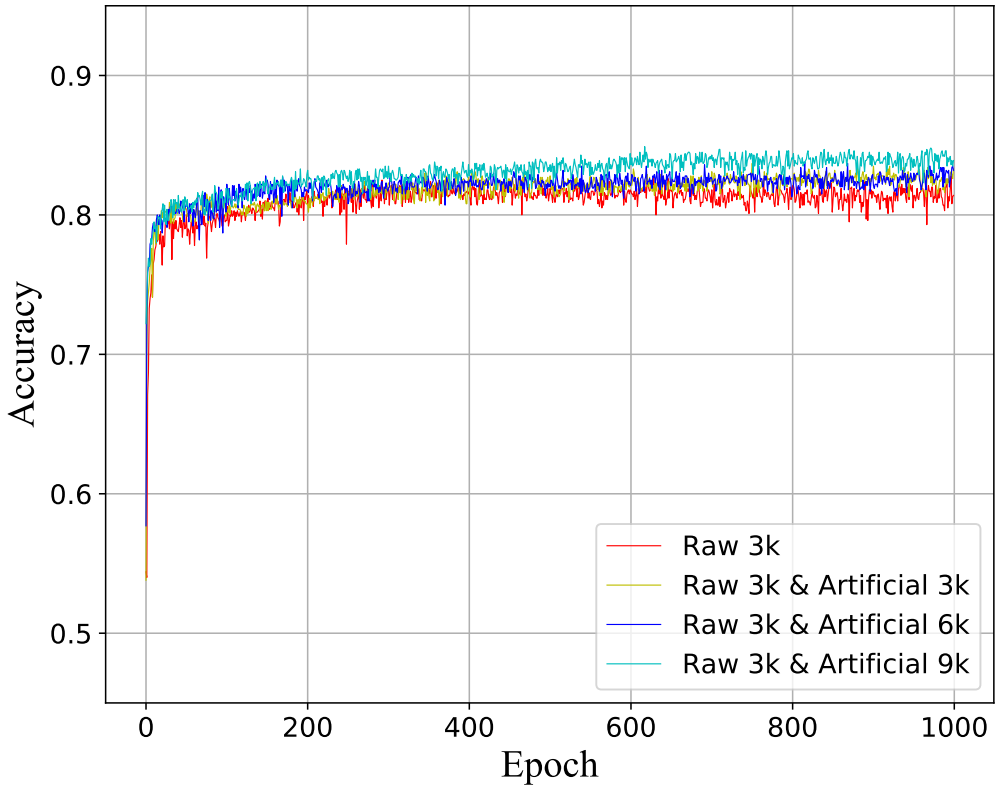

Figure 6.3: Results on different training set, test accuracy vs. number of epochs (using the 1D-CNN model).

of manual data labeling by a specialist. With this approach, supervised learning becomes more useful in medical applications such as epileptic focus localization, and opens the door for specialists to label a much smaller set of data, leaving the automatic system to generate artificial data with the adequate characteristics to train a system and generate a suitable model.

Table 6.1: Average test accuracy over the last ten epochs of 1D-CNN model with different training set.

Training set	Accuracy [%] (Mean & Std)
Raw 3k	81.52 (0.67)
Raw 3k & Artificial 3k	82.71 (0.32)
Raw 3k & Artificial 6k	82.98 (0.44)
Raw 3k & Artificial 9k	**83.91 (0.32)**

7. Conclusion and Future Work

7.1 Discussion and Contribution

In the article, in order to reduce the workload of clinical expert, a diagnostic system and several methods are proposed, In the step of extract the feature, we use two kinds of feature extraction methods, Either filter and entropy method or STFT method, we not only care about the effect of feature extraction, but also ensuring the physical interpretation and make it interpretability to clinical experts, which is the key point in clinical practice. The bandpass filters we selected are commonly used physiological frequency bands, which are interpretability to clinical experts. Because epilepsy is caused by abnormal discharge of brain cells, and entropy is a method of energy calculation, which is just suitable for measuring epileptic brain signals. Because of the diagnostic role of spikes in epilepsy, we try to use STFT for time-frequency analysis.

In the classification step, we compared several typical supervised learning methods. From the results, we can see that the more complex network model shows better performance. Although the supervised learning method show a good performance, we want to further reduce the workload of clinical experts. Thus, we introduce the PU learning method for classification. In this way, we can training a classifier only by using a small amount of labeled data (focal signal) and a large amount of unlabeled data (focal and non-focal signal), PU learning shows some advantages, but at the same time it has some shortcomings. It works well with the balancing dataset, and needs to know the proportion of the positive

data in the unlabeled data. These factors limit their application in the real world. In order to further practicalize the method, we also proposed a data enhancement method, which generates a large amount of artificial data based on a small amount of data, and uses this method to improve the performance of the model.

Some article results on the detection of epilepsy focus as shown in Table 7.1, The results in the table show that our method can achieve good performance. And we use the method of weakly supervised learning and data augmentation. The practicality of the method can be further promoted.

Table 7.1: Localization results of focal and non-focal iEEG data of published articles by using the Bern-Barcelona Dataset.

Articles	Method proposed	Accuracy in [%]
[25]	SVM & DWT	83.07
[29]	LS-SVM & EMD, Entropy	87
[17]	LS-SVM & DWT, Entropy	84
[23]	KNN & EMD-DWT, Entropy	89.4
[66]	LS-SVM & TQWT, Entropy	84.67
[32]	SVM & BEMD	86.89
	MCNN	**92.8**

7.2 Future Work

Future work focuses on two aspects. First, currently our methods are sensitive to individual differences, we need clinical experts to partially make label for each new patient. In the future we want to find a method which can across different patients. In the model, we plan use methods to analyze the common characteristics and individual differences of different patients. This way we can improve the generalization of the model for new patient data. Second, in the methods so far, one feature extraction is used each time. In the hospital's

diagnosis and treatment, doctors often use a variety of examination results to comprehensive diagnosis. It is planned to use a combination of multiple features in the future. In order to make the algorithm more practical, we need to propose targeted methods to solve the problem of clinical data imbalance.

Bibliography

[1] M Patricia and O Shafer, "About epilepsy: The basics epilepsy foundation," *http://www.epilepsy.com/learn/about-epilepsy-basics.*, 2014.

[2] Robert S Fisher, Carlos Acevedo, Alexis Arzimanoglou, Alicia Bogacz, J Helen Cross, Christian E Elger, Jerome Engel, Lars Forsgren, Jacqueline A French, Mike Glynn, et al., "ILAE official report: A practical clinical definition of epilepsy," *Epilepsia*, vol. 55, no. 4, pp. 475–482, 2014.

[3] Robert S Fisher, J Helen Cross, Carol D'souza, Jacqueline A French, Sheryl R Haut, Norimichi Higurashi, Edouard Hirsch, Floor E Jansen, Lieven Lagae, Solomon L Moshé, et al., "Instruction manual for the ILAE 2017 operational classification of seizure types," *Epilepsia*, vol. 58, no. 4, pp. 531–542, 2017.

[4] Nishant Sinha, Justin Dauwels, Marcus Kaiser, Sydney S Cash, M Brandon Westover, Yujiang Wang, and Peter N Taylor, "Predicting neurosurgical outcomes in focal epilepsy patients using computational modeling," *Brain*, vol. 140, no. 2, pp. 319–332, 2016.

[5] John S Duncan, Gavin P Winston, Matthias J Koepp, and Sebastien Ourselin, "Brain imaging in the assessment for epilepsy surgery," *The Lancet Neurology*, vol. 15, no. 4, pp. 420–433, 2016.

[6] Bilal Ahmed, Carla E Brodley, Karen E Blackmon, Ruben Kuzniecky, Gilad Barash, Chad Carlson, Brian T Quinn, Werner Doyle, Jacqueline French, Orrin Devinsky, et al.,

"Cortical feature analysis and machine learning improves detection of "MRI-negative" focal cortical dysplasia," *Epilepsy & Behavior*, vol. 48, pp. 21–28, 2015.

[7] Filipe Denaur De Moraes and Daniel Antonio Callegari, "Automated detection of interictal spikes in EEG: A literature review," *Pontifficia Universidade Católica do Rio Grande do Sul Av. Ipiranga*, 2014.

[8] Felix Rosenow and Hans Lüders, "Presurgical evaluation of epilepsy," *Brain*, vol. 124, no. 9, pp. 1683–1700, 2001.

[9] S Noachtar, C Binnie, J Ebersole, F Mauguiere, A Sakamoto, and B Westmoreland, "A glossary of terms most commonly used by clinical electroencephalographers and proposal for the report form for the EEG findings. the international federation of clinical neurophysiology.," *Electroencephalography and clinical neurophysiology. Supplement*, vol. 52, pp. 21–41, 1999.

[10] Shaun S Lodder and Michel JAM van Putten, "A self-adapting system for the automated detection of inter-ictal epileptiform discharges," *PloS one*, vol. 9, no. 1, pp. e85180, 2014.

[11] Frédéric Grouiller, Rachel C Thornton, Kristina Groening, Laurent Spinelli, John S Duncan, Karl Schaller, Michael Siniatchkin, Louis Lemieux, Margitta Seeck, Christoph M Michel, et al., "With or without spikes: localization of focal epileptic activity by simultaneous electroencephalography and functional magnetic resonance imaging," *Brain*, vol. 134, no. 10, pp. 2867–2886, 2011.

[12] J Jing, J Dauwels, T Rakthanmanon, E Keogh, SS Cash, and MB Westover, "Rapid annotation of interictal epileptiform discharges via template matching under dynamic time warping," *Journal of neuroscience methods*, vol. 274, pp. 179–190, 2016.

[13] Zhanfeng Ji, Takenao Sugi, Satoru Goto, Xingyu Wang, and Masatoshi Nakamura, "Multi-channel template extraction for automatic EEG spike detection," in *Complex*

Medical Engineering (CME), 2011 IEEE/ICME International Conference on. IEEE, 2011, pp. 179–184.

[14] K Vijayalakshmi and Appaji M Abhishek, "Spike detection in epileptic patients EEG data using template matching technique," *International Journal of Computer Applications*, vol. 2, no. 6, pp. 5–8, 2010.

[15] Jing Jin, Justin Dauwels, Sydney Cash, and M Brandon Westover, "Spikegui: Software for rapid interictal discharge annotation via template matching and online machine learning," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 2014, pp. 4435–4438.

[16] Loukianos Spyrou and Saeid Sanei, "Coupled dictionary learning for multimodal data: An application to concurrent intracranial and scalp EEG," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 2349–2353.

[17] Rajeev Sharma, Ram Bilas Pachori, and U Rajendra Acharya, "An integrated index for the identification of focal electroencephalogram signals using discrete wavelet transform and entropy measures," *Entropy*, vol. 17, no. 8, pp. 5218–5240, 2015.

[18] Andreas Antoniades, Loukianos Spyrou, David Martin-Lopez, Antonio Valentin, Gonzalo Alarcon, Saeid Sanei, and Clive Cheong Took, "Deep neural architectures for mapping scalp to intracranial EEG," *International journal of neural systems*, p. 1850009, 2018.

[19] Alexander Rosenberg Johansen, Jing Jin, Tomasz Maszczyk, Justin Dauwels, Sydney S Cash, and M Brandon Westover, "Epileptiform spike detection via convolutional neural networks," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 754–758.

[20] Saeid Sanei, *Adaptive processing of brain signals*, John Wiley & Sons, 2013.

[21] Andreas Antoniades, Loukianos Spyrou, David Martín-Lopez, A. Valentin, Gonzalo Alarcón, Saeid Sanei, and Clive Cheong Took, "Detection of interictal discharges with convolutional neural networks using discrete ordered multichannel intracranial EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, pp. 2285–2294, 2017.

[22] Danilo P Mandic and Jonathon Chambers, *Recurrent neural networks for prediction: learning algorithms, architectures and stability*, John Wiley & Sons, Inc., 2001.

[23] Anindya Bijoy Das and Mohammed Imamul Hassan Bhuiyan, "Discrimination and classification of focal and non-focal EEG signals using entropy-based features in the EMD-DWT domain," *Biomedical Signal Processing and Control*, vol. 29, pp. 11–21, 2016.

[24] S Deivasigamani, Chinnaiyan Senthilpari, and Wong Hin Yong, "Classification of focal and nonfocal EEG signals using ANFIS classifier for epilepsy detection," *International Journal of Imaging Systems and Technology*, vol. 26, no. 4, pp. 277–283, 2016.

[25] Duo Chen, Suiren Wan, and Forrest Sheng Bao, "Epileptic focus localization using EEG based on discrete wavelet transform through full-level decomposition," in *Machine Learning for Signal Processing (MLSP), 2015 IEEE 25th International Workshop on*. IEEE, 2015, pp. 1–6.

[26] Jing Zhou, Robert J Schalkoff, Brian C Dean, and Jonathan J Halford, "Morphology-based wavelet features and multiple mother wavelet strategy for spike classification in EEG signals," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 3959–3962.

[27] Yinxia Liu, Weidong Zhou, Qi Yuan, and Shuangshuang Chen, "Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG," *IEEE*

transactions on neural systems and rehabilitation engineering, vol. 20, no. 6, pp. 749–755, 2012.

[28] Samanwoy Ghosh-Dastidar, Hojjat Adeli, and Nahid Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," *IEEE transactions on biomedical engineering*, vol. 54, no. 9, pp. 1545–1551, 2007.

[29] Rajeev Sharma, Ram Bilas Pachori, and U Rajendra Acharya, "Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals," *Entropy*, vol. 17, no. 2, pp. 669–691, 2015.

[30] Forrest Sheng Bao, Jue-Ming Gao, Jing Hu, Donald YC Lie, Yuanlin Zhang, and KJ Oommen, "Automated epilepsy diagnosis using interictal scalp EEG," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*. IEEE, 2009, pp. 6603–6607.

[31] Pushpendra Singh and Ram Bilas Pachori, "Classification of focal and nonfocal EEG signals using features derived from fourier-based rhythms," *Journal of Mechanics in Medicine and Biology*, vol. 17, no. 07, pp. 1740002, 2017.

[32] Tatsunori Itakura and Toshihisa Tanaka, "Epileptic focus localization based on bivariate empirical mode decomposition and entropy," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2017*. IEEE, 2017, pp. 1426–1429.

[33] Tatsunori Itakura, Ito Shintaro, Toshihisa Tanaka, and Sugano Hidenori, "Effective frequency bands and features for epileptic focus detectionfrom interictal electrocorticogram," *TECHNICAL REPORT OF IEICE*, pp. 311–316, 2018.

[34] Xuyang Zhao, Toshihisa Tanaka, Kong. Wanzeng, Zhao. Qibin, Cao. Jianting, Sugano. Hidenori, and Yoshida. Noburu, "Epileptic focus localization based on iEEG by using

positive unlabeled (PU) learning," in *Proceedings, Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2018*, 2018, pp. 493–497.

[35] Ralph G Andrzejak, Klaus Lehnertz, Florian Mormann, Christoph Rieke, Peter David, and Christian E Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, no. 6, pp. 061907, 2001.

[36] Ralph G Andrzejak, Kaspar Schindler, and Christian Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients," *Physical Review E*, vol. 86, no. 4, pp. 046206, 2012.

[37] Vangelis P Oikonomou, Alexandros T Tzallas, and Dimitrios I Fotiadis, "A kalman filter based methodology for EEG spike enhancement," *Computer methods and programs in biomedicine*, vol. 85, no. 2, pp. 101–108, 2007.

[38] Themis P Exarchos, Alexandros T Tzallas, Dimitrios I Fotiadis, Spiros Konitsiotis, and Sotirios Giannopoulos, "EEG transient event detection and classification using association rules," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 3, pp. 451–457, 2006.

[39] Scott B Wilson and Ronald Emerson, "Spike detection: a review and comparison of algorithms," *Clinical Neurophysiology*, vol. 113, no. 12, pp. 1873–1881, 2002.

[40] Jean Gotman, "Automatic seizure detection: improvements and evaluation," *Electroencephalography and clinical Neurophysiology*, vol. 76, no. 4, pp. 317–324, 1990.

[41] Jean Gotman, "Automatic detection of seizures and spikes," *Journal of Clinical Neurophysiology*, vol. 16, no. 2, pp. 130–140, 1999.

[42] Jean Gotman, "Noninvasive methods for evaluating the localization and propagation of epileptic activity," *Epilepsia*, vol. 44, pp. 21–29, 2003.

[43] Su Liu, Candan Gurses, Zhiyi Sha, Michael M Quach, Altay Sencer, Nerses Bebek, Daniel J Curry, Sujit Prabhu, Sudhakar Tummala, Thomas R Henry, et al., "Stereotyped high-frequency oscillations discriminate seizure onset zones and critical functional cortex in focal epilepsy," *Brain*, vol. 141, no. 3, pp. 713–730, 2018.

[44] Jikai Chen and Guoqing Li, "Tsallis wavelet entropy and its application in power signal analysis," *Entropy*, vol. 16, no. 6, pp. 3009–3025, 2014.

[45] N Kannathal, Min Lim Choo, U Rajendra Acharya, and PK Sadasivan, "Entropies for detection of epilepsy in EEG," *Computer Methods and Programs in Biomedicine*, vol. 80, no. 3, pp. 187–194, 2005.

[46] Chrysostomos L Nikias and Jerry M Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, vol. 10, no. 3, pp. 10–37, 1993.

[47] Steven M Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.

[48] Joshua S Richman and J Randall Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[49] Christoph Bandt and Bernd Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical Review Letters*, vol. 88, no. 17, pp. 174102, 2002.

[50] Corinna Cortes and Vladimir Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[51] Sergey Ioffe and Christian Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[52] Bing Liu, Wee Sun Lee, Philip S Yu, and Xiaoli Li, "Partially supervised classification of text documents," *International Conference on Machine Learning*, vol. 2, pp. 387–394, 2002.

[53] Meet P Shah, SN Merchant, and Suyash P Awate, "Abnormality detection using deep neural networks with robust quasi-norm autoencoding and semi-supervised learning," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018, pp. 568–572.

[54] Zara Alaverdyan, Julien Jung, Romain Bouet, and Carole Lartizien, "Regularized siamese neural network for unsupervised outlier detection on brain multiparametric magnetic resonance imaging: application to epilepsy lesion screening," in *2018 International Conference on Medical Imaging with Deep Learning*, 2018.

[55] Zaruhi Alaverdyan, *Unsupervised representation learning for anomaly detection on neuroimaging. Application to epilepsy lesion detection on brain MRI*, Ph.D. thesis, Université de Lyon, 2019.

[56] Avrim Blum and Tom Mitchell, "Combining labeled and unlabeled data with co-training," *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pp. 92–100, 1998.

[57] Xiaoli Li and Bing Liu, "Learning to classify texts using positive and unlabeled data," *International Joint Conference on Artificial Intelligence*, pp. 587–592, 2003.

[58] Bing Liu, Yang Dai, Xiaoli Li, Wee Sun Lee, and Philip S Yu, "Building text classifiers using positive and unlabeled examples," *International Conference on Data Mining*, pp. 179–186, 2003.

[59] Wee Sun Lee and Bing Liu, "Learning with positive and unlabeled examples using weighted logistic regression," *International Conference on Machine Learning*, pp. 448–455, 2003.

[60] Charles Elkan and Keith Noto, "Learning classifiers from only positive and unlabeled data," *International Conference on Knowledge Discovery and Data Mining*, pp. 213–220, 2008.

[61] Marthinus Du Plessis, Gang Niu, and Masashi Sugiyama, "Analysis of learning from positive and unlabeled data," *Advances in Neural Information Processing Systems*, pp. 703–711, 2014.

[62] Marthinus Du Plessis, Gang Niu, and Masashi Sugiyama, "Convex formulation for learning from positive and unlabeled data," *International Conference on Machine Learning*, pp. 1386–1394, 2015.

[63] Ryuichi Kiryo, Gang Niu, Marthinus C du Plessis, and Masashi Sugiyama, "Positive-unlabeled learning with non-negative risk estimator," *Advances in Neural Information Processing Systems*, pp. 1674–1684, 2017.

[64] Josep Dinarès-Ferran, Rupert Ortner, Christoph Guger, and Jordi Solé-Casals, "A new method to generate artificial frames using the empirical mode decomposition for an EEG-based motor imagery BCI," *Frontiers in neuroscience*, vol. 12, pp. 308, 2018.

[65] Zhiwen Zhang, Feng Duan, Jordi Solé-Casals, Josep Dinarès-Ferran, Andrzej Cichocki, Zhenglu Yang, and Zhe Sun, "A novel deep learning approach with data augmentation to classify motor imagery signals," *IEEE Access*, vol. 7, pp. 15945–15954, 2019.

[66] Abhijit Bhattacharyya, Ram Pachori, Abhay Upadhyay, and U Acharya, "Tunable-Q wavelet transform based multiscale entropy measure for automated classification of epileptic EEG signals," *Applied Sciences*, vol. 7, no. 4, pp. 385, 2017.