Comparison of Handcrafted Features and Autoencoders for Epileptic Seizure Detection

Fatemeh Akrami, Fynn Aurand, Dominik T. Brockmann, Lena Faske, Laura Tiemann
University of Osnabrueck, Neurodynamics SoSe 2022
Osnabrueck, Germany
{fakrami, faurand, dobrockmann, lfaske, lautiemann}@uni-osnabrueck.de

I. Introduction

Epilepsy is one of the neurological disorders which is related to irregular functions in the central nervous system [10]. Seizure is a sudden attack that is usually associated with this disorder. It shows an abnormality in the electrical activity of the brain and can affect some parts of the body [8], [9]. Typically, the seizure can happen at any time and sometimes leads to serious injuries [11]. Furthermore, people who suffer from epileptic seizures may experience psychological burdens due to lacking proper social status [8]. Almost 50 million people are diagnosed with this disorder and the rate of its prevalence is about 0.5-10 percent [11]. So, finding an accurate diagnosis method is a crucial issue [8].

Electroencephalography (EEG) is one of the most functional neuroimaging techniques and helps us to diagnose seizures [6], [9], [10]. EEG records the brain's electrical activity as signals by using a set of electrodes [10]. These signals in the frequency domain are mostly preferred as they are inexpensive, portable, and non-invasive [8], [10]. However, the EEG signals sometimes contain some artifacts due to power supply, muscle activity, and electrode movement [9]. And these noisy EEG signals can pose a challenge to diagnose epileptic seizures. Here, artificial intelligence (AI) helps us to tackle this difficulty and also provides an opportunity to predict epileptic seizures based on EEG modalities [9]. In the field of epileptic seizures, AI techniques employ conventional methods such as machine learning (ML) and deep learning (DL) [11]. In recent years, these methods have been widely used to detect epileptic seizures by using algorithms to extract important features of the EEG signals such as auto-regressive (AR), principal component analysis (PCA), and so on [10]. But there is still a problem: EEG signals may cover irrelevant and dispensable features, as the recorded activity usually consists of multi-channel signals [8]. Therefore, the accuracy of the classification of the features may be reduced. Now, there is a question: How can we improve the quality and reliability of the output (seizure prediction and detection) by choosing the best-suited features in the classification process?

The classification process detects the similarities among groups and classes. Fig 1 presents a schematic view of this process [10]. We can apply this approach to distinguish between seizure and non-seizure signals. Some algorithms like artificial neural networks (ANN), support vector machines (SVM), and decision trees are known as classifier [10]. In this project, we will mainly focus on which features of the data are best suited for classification purposes.

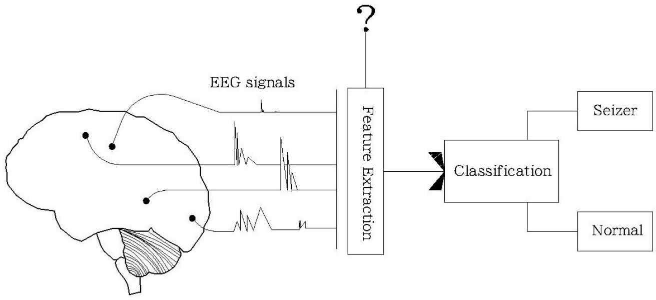

Fig. 1. This scheme simply shows the process of classification. After recording the EEG signal, the feature extraction is used to put the signals in the appropriate classes, so the seizure signals are detectable.

II. Theoretical background

An automatic seizure detection system consists of many steps: First of all, EEG data needs to be acquired and preprocessed. However, the two main stages of such a system are: feature extraction and classification [13]. During the feature extraction step, the most representative or characterizing features of the EEG data need to be extracted and selected [3]. This step plays an immense role: it directly affects the precision and sophistication of the classifier, as the classification system is built and trained upon these extracted features for the discrimination of different EEG signal types. [2], [3], [13]

Strategies for extracting features can be categorized as (1) automatic feature learning from signals or (2) handcrafting features [7]. Most of the work uses handcrafted features [13]. These features can be divided into the time domain, frequency domain, time-frequency domain and non-linear domain [7].

Various methods for handcrafted feature extraction for EEG seizure detection have been proposed: For example, Fourier transform and wavelet transform are used to extract frequency-domain features form EEG signals and their results can be used for EEG seizure detection [13]. Handcrafted features from the time domain would be for example the approximate entropy, sample entropy or other statistical measures [2]. Also, there has been an growing interest in introducing time-frequency image descriptors to EEG seizure detection. [13]

However, the design of these handcrafted features often involves a significant amount of expert knowledge to derive the characteristic epileptic patterns in the EEG signals. Additionally, it isn't generally ensured that these manually extracted features are optimal. Therefore, extremely good efforts have been exerted to enhance automatic seizures detection systems based on EEG signals' recordings [3]. It has been proven, that automatically learned features have been more robust than handcrafted features and that they achieve better detection performance [13].

Thus, we want to compare automatic seizure detection systems using hand-crafted features against systems using automatic feature learning.

III. Methodology

Dataset. There exist multiple EEG datasets that contain labeled data to train systems for automatic epileptic seizure detection. The dataset provided by the University of Bonn [4] is observed to be widely used. This dataset consists of EEG data of 10 patients obtained intracranially (IEEG) and from the scalp. The data is labeled with two classes: healthy and epileptic patients [9]. As the Bonn dataset is widely employed and publicly available, we will use it in our study.

Feature Extraction. We will compare the performance of a classification system trained using hand-picked features to features obtained by autoencoders. Autoencoders can reduce the dimensionality of data by using an encoding/decoding pattern. It uses the input as the target output and is thereby an unsupervised learning algorithm. By reducing the dimensions, features of the data can be extracted that are important for the classification regarding a specific task. [5]

Classification. In order to classify the data using the obtained feature vector as input, we will use a softmax classifier. A softmax classifier is an interconnected artificial neural network that uses the logistic sigmoid function as the activation function. The output of this classifier is a probability vector for each class. In the case of a two class problem, the class with a value greater than 0.5 is selected. [12]

Evaluation. To compare the two feature extraction methods, we will evaluate the performance by comparing accuracy, sensitivity and specificity of the two resulting classification systems. The accuracy can show the overall performance, while the latter two are especially important to determine the ability to classify diseased and healthy cases accurately [12]. Additionally, we may compare the time, effort and computational complexity needed for each method in relation to the resulting performance.

Implementation. The implementation of preprocessing, feature extraction and classification will be done using Python. The machine learning models will be implemented using Tensorflow [1]. The full source code will be made public using a public GitHub repository.

IV. Time Plan

Week	Task	Responsibility
23.-29.05.	Identifying Features	Fatemeh, Laura
30.-05.06.	Feature Extraction	Fynn, Lena
06.-12.06.	Model Development	Dominik
13.-19.06.	Evaluation	All
20.-03.07.	Poster Creation	All
04.-10.07.	First Draft of Paper	All
11.-17.07.	Final Draft of Paper	All
18.-25.07.	Final Paper Adjustments	All

Since our group has different strengths and weaknesses, we decided to divide the main tasks under those whose skillset fit best to the specific tasks and arising problems. Therefore, Fatemeh and Laura will focus on the literature, the important features and time management, while Fynn and Lena will cover the feature extraction. Since Dominik has further experience with Machine Learning, his responsibility lies within the development of the model. Parts like evaluation, poster presentation and the writing of the paper will be done by every team member and divided in a reasonable manner.

References

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[2] Ahmed Abdelhameed and Magdy Bayoumi. A deep learning approach for automatic seizure detection in children with epilepsy. *Frontiers in Computational Neuroscience*, 15:29, 2021.

[3] Ahmed M Abdelhameed, Hisham G Daoud, and Magdy Bayoumi. Epileptic seizure detection using deep convolutional autoencoder. In *2018 IEEE International Workshop on Signal Processing Systems (SiPS)*, pages 223–228. IEEE, 2018.

[4] 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. *Phys. Rev. E*, 64:061907, Nov 2001.

[5] Qinxue Meng, Daniel Catchpoole, David Skillicom, and Paul J. Kennedy. Relational autoencoder for feature extraction. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 364–371, 2017.

[6] Marzieh Savadkoohi, Timothy Oladunni, and Lara A. Thompson. A machine learning approach to epileptic seizure prediction using electroencephalogram (eeg) signal. *Biocybernetics and Biomedical Engineering*, 2020.

[7] Afshin Shoeibi, Navid Ghassemi, Roohallah Alizadehsani, Modjtaba Rouhani, Hossein Hosseini-Nejad, Abbas Khosravi, Maryam Panahiazar, and Saeid Nahavandi. A comprehensive comparison of handcrafted features and convolutional autoencoders for epileptic seizures detection in eeg signals. *Expert Systems with Applications*, 163:113788, 2021.

[8] Afshin Shoeibi, Navid Ghassemi, Marjane Khodatars, Mahboobeh Jafari, Parisa Moridian, Roohallah Alizadehsani, Ali Khadem, Yinan Kong, Assef Zare, Juan Manuel Górriz, Javier Ramírez, Maryam Panahiazar, Abbas Khosravi, and Saeid Nahavandi. Applications of epileptic seizures detection in neuroimaging modalities using deep learning techniques: Methods, challenges, and future works. *arXiv: Learning*, 2021.

[9] Afshin Shoeibi, Marjane Khodatars, Navid Ghassemi, Mahboobeh Jafari, Parisa Moridian, Roohallah Alizadehsani, Maryam Panahiazar, Fahime Khozeimeh, Assef Zare, Hossein Hosseini-Nejad, Abbas Khosravi, Amir F. Atiya, Diba Aminshahidi, Sadiq Hussain, Modjtaba Rouhani, Saeid Nahavandi, and Udyavara Rajendra Acharya. Epileptic seizures detection using deep learning techniques: A review. *International Journal of Environmental Research and Public Health*, 18(11), 2021.

[10] Athar A. Ein Shoka, Monagi H Alkinani, Monagi H. Alkinani, A. S. El-Sherbeny, Ayman El-Sayed, Ayman El-Sayed, and Mohamed M. Dessouky. Automated seizure diagnosis system based on feature extraction and channel selection using eeg signals. *Brain Informatics*, 2021.

[11] Mohammad Khubeb Siddiqui, Ruben Morales-Menendez, Xiaodi Huang, Xiaodi Huang, and Nasir Hussain. A review of epileptic seizure detection using machine learning classifiers. *Brain Informatics*, 2020.

[12] Sharathappriyaa Venkatesan, S Gautham, and R Lavanya. Auto-encoder based automated epilepsy diagnosis. pages 976–982, 09 2018.

[13] Ye Yuan, Guangxu Xun, Kebin Jia, and Aidong Zhang. A multi-view deep learning framework for eeg seizure detection. *IEEE journal of biomedical and health informatics*, 23(1):83–94, 2018.