Received: 15 January 2022 | Revised: 11 April 2022 | Accepted: 11 April 2022

DOI: 10.1111/epi.17259

Epilepsia™

SPECIAL ISSUE ARTICLE

Seizure detection with deep neural networks for review of two-channel electroencephalogram

Manfred Hartmann[1] | Johannes Koren[2,3] | Christoph Baumgartner[2,3,4] | Jonas Duun-Henriksen[5] | Gerhard Gritsch[1] | Tilmann Kluge[1] | Hannes Perko[1] | Franz Fürbass[1]

[1]Center for Health and Bioresources, Austrian Institute of Technology, Vienna, Austria

[2]Department of Neurology, Clinic Hietzing, Vienna, Austria

[3]Karl Landsteiner Institute for Clinical Epilepsy Research and Cognitive Neurology, Vienna, Austria

[4]Medical Faculty, Sigmund Freud University, Vienna, Austria

[5]UNEEG medical, Allerød, Denmark

Correspondence
Manfred Hartmann, Giefinggasse 4, 1210 Vienna, Austria.
Email: manfred.hartmann@ait.ac.at

Abstract

Ultra-long-term electroencephalographic (EEG) registration using minimally invasive low-channel devices is an emerging technology to assess sporadic seizure events. Highly sensitive automatic seizure detection algorithms are needed for semiautomatic evaluation of these prolonged recordings. We describe the design and validation of a deep neural network for two-channel seizure detection. The model is trained using EEG recordings from 590 patients in a publicly available seizure database. These recordings are based on the full 10–20 electrode system and include seizure annotations created by reviews of the full set of EEG channels. Validation was performed using 48 scalp EEG recordings from an independent epilepsy center and consensus seizure annotations from three neurologists. For each patient, a three-electrode subgroup (two channels with a common reference) of the full montage was selected for validation of the two-channel model. Mean sensitivity across patients of 88.8% and false positive rate across patients of 12.9/day were achieved. The proposed training approach is of great practical relevance, because true recordings from low-channel devices are currently available only in small numbers, and the generation of gold standard seizure annotations in two EEG channels is often difficult. The study demonstrates that automatic seizure detection based on two-channel EEG data is feasible and review of ultra-long-term recordings can be made efficient and effective.

KEYWORDS

deep learning, seizure detection, subscalp EEG, ultra-long-term EEG

1 | INTRODUCTION

With recent advances in wearable devices and signal analysis, the field of mobile health has undergone a remarkable evolution. Regarding epilepsy, numerous medical devices are available today that record electroencephalograms (EEGs) on an outpatient basis, at home, and during the patient's daily life and activities. In particular, minimally invasive, subscalp implanted low-channel EEG devices are an emerging technology enabling ultra-long-term EEG registration to assess sporadic and rare seizure events and to track seizure burden over months or years. The number of electrodes in these devices is often limited, leading to a reduced spatial coverage of brain regions. Identification of epileptic seizure events in ultra-long-term EEG recordings for diagnostic purposes and therapy control leads to a massive review burden due to the vast amount of data recorded by these devices. This creates a strong demand for highly sensitive automatic seizure detection algorithms, which can be utilized in semiautomatic assessments of these prolonged recordings.

Interpretation of EEGs in a clinical context is largely based on visual interpretation of at least 19 channels based on the International Federation of Clinical Neurophysiology electrode array,[1] making detection of seizures in two-channel recordings difficult even for expert human readers (cf. Figure 2, with four examples of two-channel EEGs). The limited spatial coverage leads to poorly covered brain regions, which can result in limited sensitivity for focal seizures. Even for the experienced human reader, interpretation of such EEGs is difficult because of the limited ability to compare spatially contiguous EEG channels to detect artifacts and assess spatial differences across brain regions. Computational models for automatic detection of seizures in EEG recordings based on few electrodes face similar challenges and have been tackled in recent literature, for example, for seizure detection in recordings from electrodes behind the ear,[2–5] or from wireless, wearable EEG sensors.[6] Very recently, a large validation study on automatic detection of absences using a wearable headband with dry electrodes in 102 patients was reported,[7] where average sensitivity per patient was 78.8%, with.53 false positives (FPs) per hour on average, for this specific seizure type. In the present work, we address devices with two channels, that is, three electrodes including the reference, arranged approximately on a straight line and with contact spacings similar to the 10–20 electrode array. We demonstrate the possibility of detecting seizures with high accuracy using deep neural networks even with only three EEG electrodes. The proposed approach enables end-to-end training, that is, input of unprocessed EEG waveforms, which promises to achieve high accuracy and great flexibility in adapting to varying conditions.

Key Points

- A deep neural network for two-channel seizure detection is designed and trained using EEG based on the full 10–20 electrode system
- Mean sensitivity across patients of 88.8% and low false positive rate across patients of 12.9 per day are achieved in a validation study including 48 patients from an independent center
- The study demonstrates that review of ultra-long-term two-channel EEG recordings can be made efficient and effective using automatic seizure detection

2 | MATERIALS AND METHODS

2.1 | Model architecture

To achieve a high degree of accuracy, the neural network architecture must be able to extract EEG features that characterize epileptic seizures: first, the morphology of the EEG waveforms, the "primary feature" of pathological epileptiform EEG activity; and second, the spatial field of the EEG potentials to allow discrimination of artifacts. However, the possibilities for exploiting this feature are limited when only two channels are available. Third, for accurate detection of electrographic seizure patterns, the temporal evolution of EEG parameters is particularly important. The progression of amplitude, frequency, and morphology from the preictal state through seizure onset to the development in the ictal course are characteristic features in the EEG of a seizure. Based on these assumptions, we developed the following neural network architecture.

At the input, artifacts are reduced using the PureEEG algorithm,[8] whose parameters are tuned for operation with only two EEG channels. Mainly for the extraction of morphological, but also for spatial features, the model then continues with a deep stack of convolutional layers structured as a residual neural network. The features extracted by these layers have different abstraction levels and time scales. In the next block, these features are combined by a bidirectional feature pyramid network[9] that allows evaluation of temporally local features in a more global context. For an EEG model, this may mean that, for example, epileptiform discharges are detected and simultaneously evaluated in terms of their occurrence against an interictal background and their evolution in amplitude, frequency, and morphology. In the next block, onset position estimation network generates estimates for potential seizure onset times, which is inspired by object recognition models such as *Faster R-CNN* for image processing[10] and interictal discharges in EEG.[11] The final block in the model, a multilayer perceptron classifier, is time-aligned according to the estimated seizure onset time and distinguishes ictal EEG from interictal EEG.

2.2 | Training

In addition to a suitable architecture, it is equally important for the successful development of an artificial neural network to use training data in sufficient quantity and quality. Large collections of 10–20-based EEG recordings with seizure labels that can be used for this purpose exist in various publicly available databases. The model for this study was trained with data from the Temple University Hospital Seizure Detection Corpus.[12] Labels therein are based on clinical reports in combination with visual review of the full 10–20 channel set EEG waveforms by a team of highly trained students. This database is a collection of 1184 EEG recordings from 592 patients from different departments, including epilepsy monitoring and intensive care units. It includes 2377 seizure annotations with time points of electrographic seizure onset and offset and with seizure types, which include 1868 (79%) focal seizures (with and without impaired awareness) and 509 (21%) generalized seizures (generalized tonic–clonic seizures, myoclonic seizures, and absences).

To train our two-channel model with 10–20 montage recordings, we defined 14 two-channel subgroups of the full 10–20 electrode set. This was done so that electrodes within each subgroup are approximately on straight lines on the scalp, as illustrated in Figure 1. The subgroups were also distributed to fully cover the entire 10–20 electrode set, so that any seizure visible in the full 10–20 channel set should also be visible in at least one subgroup.

For the training task, 14 separate models shared their parameters but independently processed data from one of each of the 14 subgroups. As a result, in (focal) seizure samples, some subgroups contain data without any visible seizure patterns. A training loss function that updates the model parameters based on the electrodes with the most pronounced seizure pattern was designed and optimized. This ensures that the model can be applied to a single pair of channels and still produce correct results.

2.3 | Validation

For the validation study, we retrospectively included scalp EEG recordings of 50 subjects who underwent video-EEG monitoring in an epilepsy monitoring unit in Vienna (Clinic Hietzing, Rosenhügel) for the purpose of differential diagnosis or presurgical evaluation in the years 2009 and 2012. We included the first available 29 subjects recorded in 2009 and the first available 21 subjects recorded in 2012 who were 18 years of age or older and experienced seizure events during the recording. All were subsequently diagnosed with epilepsy. Two subjects were excluded because of an unusual abundance of interictal findings, leading to difficulties in separating ictal and interictal EEG to establish a meaningful gold standard for electrographic seizures. Mean age of the remaining 48 patients during recording was 38.5 years (minimum = 18, maximum = 76), 45 patients (94%) were suffering from focal epilepsy (26 patients [54%] with temporal lobe epilepsy, 19 patients [40%] with extratemporal lobe or multifocal epilepsy), and three patients (6%) had generalized or unknown epilepsy. Types of seizure in the validation data comprised focal seizures with and without awareness, focal-to-bilateral tonic–clonic seizures, and generalized tonic–clonic seizures. Validation was done as follows. The first 30-h period for each patient containing at least one seizure (according to the original information in the

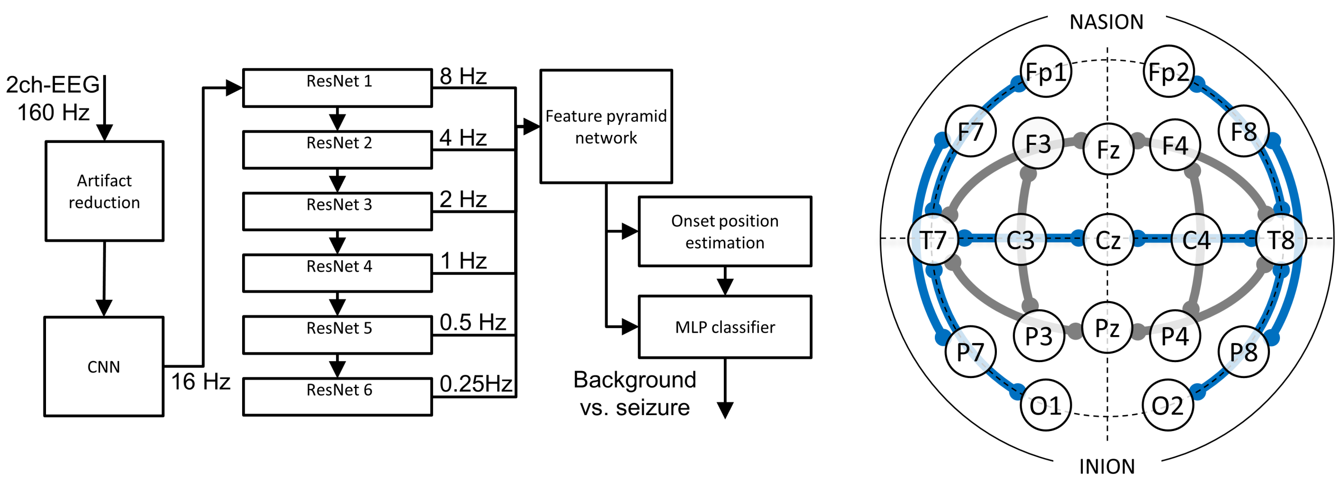

FIGURE 1 Neural network design. Left: Block diagram of the neural network architecture. Right: Arrangement of 14 electrode triplets used to train the two-channel detectors indicated by blue and gray lines. Electrode-triplets (including the reference electrodes) used for validation, which were selected by a reviewer for each individual patient, are indicated by blue lines. CNN, convolutional neural network; EEG, electroencephalogram; ResNet, residual network; MLP, multilayer perceptron

clinical report) was selected for review. This was done to unify the amount of data from each patient and to limit the review effort and led to an inclusion of 121 seizures in total. Data were then blindly reviewed by three independent neurologists. Majority voting was used to create the ground truth seizure annotations, that is, in absence of full consensus, decisions were taken from two of three reviewers. Next, based on the clinical report and review of ictal EEGs, one reviewer (F.F.) assigned an electrode subset representing a "virtual two-channel device" for each patient, where seizures would most likely be captured. This electrode subset was used to detect individual patients' seizures. The reviewer could select one of the eight subgroups marked in blue in Figure 1. For evaluation of the model performance on each patient, only detections from the selected electrode subgroup were considered. Seizure epochs were counted as true positive (TP) if at least one detection occurred within the consensus annotation time range or within 40 s before the consensus annotation time range, and as false negative (FN) otherwise. Any detection outside of an extended consensus annotation was counted as FP. A 40-s extension window before consensus annotation was included, because the detection model was trained to detect the seizure onset time point, whereas the reviewers' consensus annotations marked the clear-cut seizure patterns within the seizure and did not necessarily include the onset time of the seizures.

Performance of the model was evaluated on the patient level in terms of sensitivity = TP/(TP + FN) and FP rate (FPR) = FP/T_{EEG}, where T_{EEG} is the duration of the EEG recording.

3 | RESULTS

In the validation dataset, the deep neural network model detected all annotated seizures in 37 of 43 patients with at least one consensus seizure annotation (86%). Five patients did not have any consensus seizure annotations. Mean sensitivity across patients was 88.8% (95% confidence interval [CI] = 76.7 – 95.3), with CI determined by bootstrapping (n = 10 000). The mean FPR (per day) across patients was 12.9/day (95% CI = 8.9–19.4). The median FPR was 6.5/day, and for 41 of 48 patients we had less that one false detection per hour. Figure 2 shows scatter plots illustrating results for sensitivity and FPR, and four examples of two-channel EEGs. The first EEG shows a correctly detected seizure (TP) from a patient with focal epilepsy. The second EEG is a seizure with 13-Hz rhythmic activity that is unequivocally distinguishable from muscle artifacts with the full set of electrodes from another focal epilepsy patient missed by the algorithm (FN). The third EEG shows a rhythmic, patient-induced artifact from a bilateral temporal lobe epilepsy patient, which led to an FP detection. The fourth EEG is a TP detection of a seizure with high-amplitude repetitive spikes from a frontal lobe epilepsy patient.

4 | DISCUSSION

4.1 | Algorithm

In this work, we successfully designed a deep neural network architecture for automatic detection of seizures in two-channel EEG recordings. We showed the feasibility of training a two-channel detector with 10–20-electrode array data, wherein seizure annotations do not contain information about the spatial spread of seizure patterns. This was done by appropriate definition of 14 electrode triplets and processing of corresponding EEG data by 14 independent two-channel detectors. The outputs of all 14 detectors were evaluated separately during training but combined into a training loss function in such a way that only EEG channels in the seizure focus preferentially led to an update of the model parameters. This approach enables end-to-end training of a two-channel model using 10–20 training data without spatial information in the seizure labels. Advantages of this approach are flexibility in adapting to varying conditions, high accuracy, and computational efficiency. End-to-end training was also reported for detection of absence seizures[7] and focal seizures.[5] Due to the generalized activity in absence seizures, reduction of channels in training data from 10–20 montage annotations is trivial. For the case of focal seizures, a method for annotation correction prior to training was proposed.[5] In contrast, our approach, which follows the same goal, allows training the model without any modification of training data.

The validation and training data used in this study came from different, independent hospitals. Such a strict data regime is paramount for machine learning models to prove their generalization capabilities. We want to point out that our results are achieved with training data from only one publicly available EEG database. We hypothesize that the inclusion of additional, independent data for training—possibly recorded with the target device—would have a positive effect on performance, robustness, and generalization capability.

4.2 | Application

The goal of this work is to make review of ultra-long-term records efficient and effective, which corresponds

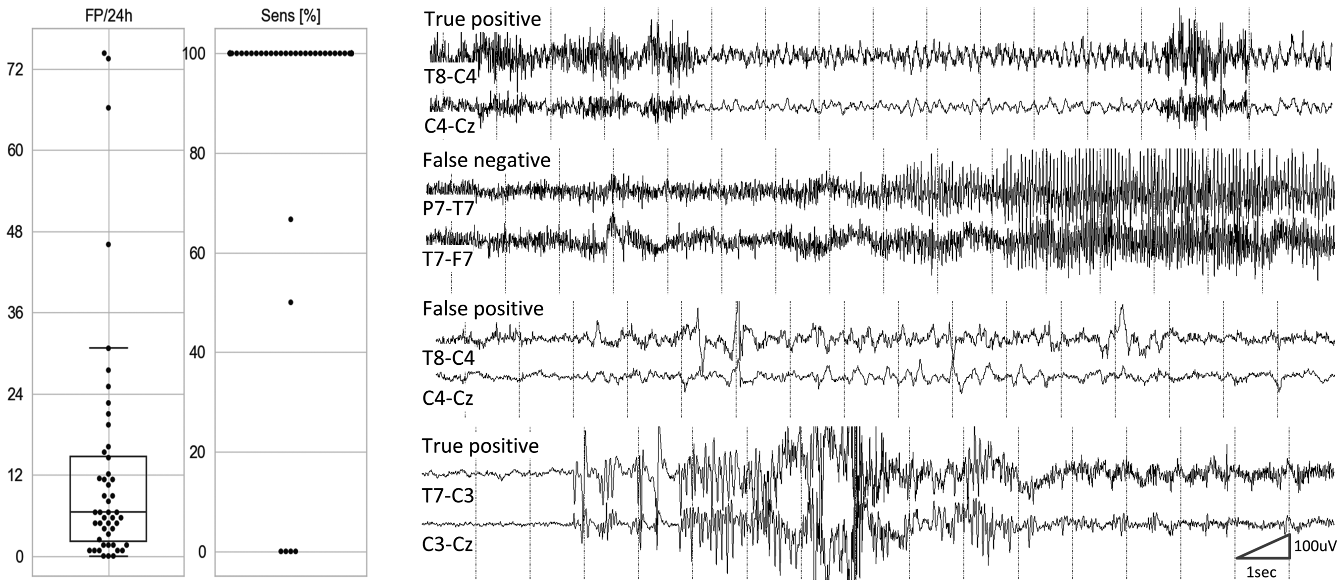

FIGURE 2 Results of automatic seizure detection. Left: Scatter plot of false positive (FP) detection rates of 48 patients and sensitivities (Sens) of 43 patients with consensus seizure annotations. Five patients did not have any consensus seizure annotations. For 41 of 48 patients, false detection rate is <1/h. Sensitivities are at 100% for 37 of 43 patients. Right: Two-channel electroencephalogram examples. The first example (true positive [TP]) and second examples (false negative) are from two focal epilepsy patients, the third example (FP) is from a bitemporal lobe epilepsy patient, and the last example (TP) is from a frontal lobe epilepsy patient

to finding the most seizures in a dataset in a short time. In the final validation dataset, which included EEG recordings from 48 patients, our neural network achieved a mean sensitivity across patients of 88.8% and a mean FPR across patients of 12.9/day. We believe that review with a sensitivity close to 90% can be considered effective, although it depends on the specific application. This value is close to published studies on non-patient-specific, medical device seizure detection software.[13,14]

A review process that is solely based on automatic detections would include assessment of approximately 400 FP detections in 1 month of continuous EEG recording on average. Therefore, review assisted by automatic detections within our approach seems to be highly efficient, considering that 1 month of data comprises approximately 86 000 EEG pages of 30 s each. As such, our results suggest the application of our model in the review of ultra-long term EEG data from mobile EEG recording devices in a semiautomated workflow. Mobile EEG recorders such as the subscalp device from UNEEG medical[15] were shown to be equivalent to scalp EEG. Therefore, our approach is suitable for continuous assessment of electrographic seizures in patients implanted with such recording devices. A limitation of our study is that the validation was done using scalp EEG recordings, assuming their equivalence to subscalp EEG recordings. Validation of our model on real subscalp recordings is ongoing.

We conclude that automatic seizure detection based on two-channel EEG data is feasible. High sensitivity as well as low false detection rate can be achieved, assuming optimum placement of the electrodes in a specific patient.

ACKNOWLEDGMENT

We address our thanks to the members of the Neural Engineering Data Consortium for creating and providing the Temple University Hospital Seizure Detection Corpus.

CONFLICT OF INTEREST

J.D.-H. is employed by UNEEG medical, which manufactures a subscalp EEG monitor. The remaining authors report no conflict of interest regarding the content of this study. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

ORCID

Manfred Hartmann https://orcid.org/0000-0002-1446-7600
Johannes Koren https://orcid.org/0000-0002-5290-2837
Jonas Duun-Henriksen https://orcid.org/0000-0003-1558-8225
Franz Fürbass https://orcid.org/0000-0002-6744-2802

REFERENCES

1. Seeck M, Koessler L, Bast T, Leijten F, Michel C, Baumgartner C, et al. The standardized EEG electrode array of the IFCN. Clin Neurophysiol. 2017;128:2070–7.

2. You S, Hwan Cho B, Shon Y-M, Seo D-W, Kim IY. Semi-supervised automatic seizure detection using personalized anomaly detecting variational autoencoder with behind-the-ear EEG. Comput Methods Programs Biomed. 2022;213:106542.
3. Vandecasteele K, De Cooman T, Dan J, Cleeren E, Van Huffel S, Hunyadi B, et al. Visual seizure annotation and automated seizure detection using behind-the-ear electroencephalographic channels. Epilepsia. 2020;61(4):766–75.
4. Swinnen L, Chatzichristos C, Jansen K, Lagae L, Depondt C, Seynaeve L, et al. Accurate detection of typical absence seizures in adults and children using a two-channel electroencephalographic wearable behind the ears. Epilepsia. 2021;62(11):2741–52.
5. Zhang J, Chatzichristos C, Vandecasteele K, Swinnen L, Broux V, Cleeren E, et al. Automatic annotation correction for wearable EEG based epileptic seizure detection. J Neural Eng. 2022;19(1):016038. https://doi.org/10.1088/1741-2552/ac54c1
6. Frankel MA, Lehmkuhle MJ, Spitz MC, Newman BJ, Richards SV, Arain AM. Wearable reduced-channel EEG system for remote seizure monitoring. Front Neurol. 2021;12:1842. https://doi.org/10.3389/fneur.2021.728484
7. Japaridze G, Loeckx D, Buckinx T, Armand Larsen S, Proost R, Jansen K, et al. Automated detection of absence seizures using a wearable electroencephalographic device: a phase-3 validation study and feasibility of automated behavioral testing. Epilepsia. 2022; 1–7. https://doi.org/10.1111/epi.17200
8. Hartmann MM, Schindler K, Gebbink TA, Gritsch G, Kluge T. PureEEG: automatic EEG artifact removal for epilepsy monitoring. Neurophysiol Clin. 2014;44(5):479–90.
9. Tan M, Pang R, Le QV. EfficientDet: Scalable and efficient object detection. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020. p. 10778–87.
10. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2017;39(6):1137–49.
11. Fürbass F, Kural MA, Gritsch G, Hartmann M, Kluge T, Beniczky S. An artificial intelligence-based EEG algorithm for detection of epileptiform EEG discharges: validation against the diagnostic gold standard. Clin Neurophysiol. 2020;131(6):1174–9.
12. Shah V, vonWeltin E, Lopez S, McHugh JR, Veloso L, Golmohammadi M, et al. The Temple University Hospital seizure detection corpus. Front Neuroinform. 2018;12:83.
13. Koren J, Hafner S, Feigl M, Baumgartner C. Systematic analysis and comparison of commercial seizure-detection software. Epilepsia. 2021;62(2):426–38.
14. Reus EEM, Visser GH, vanDijk JG, Cox FME. Automated seizure detection in an EMU setting: are software packages ready for implementation?Seizure. 2022;96:13–7.
15. Duun-Henriksen J, Baud M, Richardson MP, Cook M, Kouvas G, Heasman JM, et al. A new era in electroencephalographic monitoring? Subscalp devices for ultra–long-term recordings. Epilepsia. 2020;61(9):1805–17.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Hartmann M, Koren J, Baumgartner C, Duun-Henriksen J, Gritsch G, Kluge T, et al. Seizure detection with deep neural networks for review of two-channel electroencephalogram. Epilepsia. 2023;64:S34–S39. https://doi.org/10.1111/epi.17259