



OPEN ACCESS

EDITED BY

Giovanni Pellegrino,
McGill University Health Centre,
Canada

REVIEWED BY

Ailiang Miao,
Nanjing Brain Hospital Affiliated to
Nanjing Medical University, China
Stavros I. Dimitriadis,
Integrative Neuroimaging Lab, Greece
Pradeep Modur,
University of Texas at Austin,
United States

*CORRESPONDENCE

Xueping Peng
xueping.peng@uts.edu.au
Ting Wu
lwu123@163.com
Jiayang Guo
guojy@xmu.edu.cn
Yijie Pan
pyj20@mails.tsinghua.edu.cn

RECEIVED 07 September 2021

ACCEPTED 03 August 2022

PUBLISHED 09 September 2022

CITATION

Zhao X, Peng X, Niu K, Li H, He L,
Yang F, Wu T, Chen D, Zhang Q,
Ouyang M, Guo J and Pan Y (2022) A
multi-head self-attention deep
learning approach for detection and
recommendation of neuromagnetic
high frequency oscillations in epilepsy.
Front. Neuroinform. 16:771965.
doi: 10.3389/fninf.2022.771965

COPYRIGHT

© 2022 Zhao, Peng, Niu, Li, He, Yang,
Wu, Chen, Zhang, Ouyang, Guo and
Pan. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

A multi-head self-attention deep learning approach for detection and recommendation of neuromagnetic high frequency oscillations in epilepsy

Xiangyu Zhao^{1,2}, Xueping Peng^{3*}, Ke Niu⁴, Hailong Li⁵,
Lili He⁵, Feng Yang¹, Ting Wu^{6,7*}, Duo Chen⁸, Qiusi Zhang¹,
Menglin Ouyang⁹, Jiayang Guo^{10,11*} and Yijie Pan^{12,13*}

¹Information Technology Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, ²National Engineering Research Center for Information Technology in Agriculture, Beijing, China, ³Australian Artificial Intelligence Institute, Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW, Australia, ⁴Computer School, Beijing Information Science and Technology University, Beijing, China, ⁵Department of Radiology, Imaging Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States, ⁶Department of Radiology, Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China, ⁷Department of Magnetoencephalography, Nanjing Brain Hospital, Affiliated to Nanjing Medical University, Nanjing, China, ⁸School of Artificial Intelligence and Information Technology, Nanjing University of Chinese Medicine, Nanjing, China, ⁹The Affiliated Hospital of Medical School, Ningbo University, Ningbo, China, ¹⁰National Institute for Data Science in Health and Medicine, Xiamen University, Xiamen, China, ¹¹Department of Hematology, School of Medicine, Xiamen University, Xiamen, China, ¹²Department of Computer Science and Technology, Tsinghua University, Beijing, China, ¹³Ningbo Institute of Information Technology Application, Chinese Academy of Sciences, Ningbo, China

Magnetoencephalography is a noninvasive neuromagnetic technology to record epileptic activities for the pre-operative localization of epileptogenic zones, which has received increasing attention in the diagnosis and surgery of epilepsy. As reported by recent studies, pathological high frequency oscillations (HFOs), when utilized as a biomarker to localize the epileptogenic zones, result in a significant reduction in seizure frequency, even seizure elimination in around 80% of cases. Thus, objective, rapid, and automatic detection and recommendation of HFOs are highly desirable for clinicians to alleviate the burden of reviewing a large amount of MEG data from a given patient. Despite the advantage, the performance of existing HFOs rarely satisfies the clinical requirement. Consequently, no HFOs have been successfully applied to real clinical applications so far. In this work, we propose a multi-head self-attention-based detector for recommendation, termed MSADR, to detect and recommend HFO signals. Taking advantage of the state-of-the-art multi-head self-attention mechanism in deep learning, the proposed MSADR achieves a more superior accuracy of 88.6% than peer machine learning models in both detection and recommendation tasks. In addition, the robustness of MSADR is also extensively assessed with various ablation tests, results of which further demonstrate the effectiveness and generalizability of the proposed approach.

KEYWORDS

high frequency oscillations (HFOs), magnetoencephalography, MEG, deep learning, multi-head self-attention, HFOs detection, HFOs recommendation

1. Introduction

About 30% of pediatric patients with epilepsy are medically intractable and require respective neurosurgery to gain seizure freedom (Durnford et al., 2011; Yamakawa et al., 2020). Recording epileptic activities are crucial to the pre-operative localization of epileptogenic zones and the optimization of the diagnosis of epilepsy. The success of epilepsy surgery depends on the pre-operative localization of epileptogenic zones (Guo et al., 2018). Although intracranial electroencephalography (iEEG) is commonly treated as the gold standard for the localization of epileptogenic zones, it may bring a risk of infection and bleeding during implantation (Hu et al., 2015). Thus, a noninvasive detection method for epileptogenic zones is preferred to epilepsy surgery. Magnetoencephalography (MEG) is a noninvasive technology for the detection of epileptic activities. MEG has a higher spatial resolution to localize epileptic activities for epilepsy surgery than other noninvasive approaches, such as electroencephalography (EEG) (Nakasato et al., 1994).

Localizing epileptogenic zones play a central role in epilepsy surgery. However, to date, there are no robust biomarkers that are able to accurately capture the location of epileptogenic zones (Tamilia et al., 2017). A variety of diagnostic indicators are introduced in the current clinical practice to estimate the epileptogenic zones. Despite the progress, existing methods, which heavily rely on epileptic spikes (typically ≤ 70 Hz), fail to reduce seizure frequency in approximately 50% of the cases, which greatly limits their applications in epilepsy surgery (Stigsdotter-Broman et al., 2014; Olan Çocuklarda and Öncesi, 2015; Oldham et al., 2015; Reinholdson et al., 2015; Verdinelli et al., 2015). High frequency oscillations (HFOs) (typically 80–500 Hz) can be used to localize the epileptogenic zones as biomarkers. Recent studies (Xiang et al., 2010; Ontario, 2012; Modur, 2014; Van Klink et al., 2014; Van't Klooster et al., 2015; Leung et al., 2018; Nevalainen et al., 2020) show that applying pathological HFOs to localize the epileptogenic zones leads to a significant reduction in seizure frequency, even seizure elimination in about 80% of cases. Thus, pathological HFOs have been associated with epileptogenic zones (Xiang et al., 2009; Miao et al., 2014). There is increasing evidence to show that HFOs are putative biomarkers to identify epileptic regions of the brain, which may improve the surgical diagnosis and surgical outcomes of patients with epilepsy.

Recent reports (Papadelis et al., 2009, 2016; Van Klink et al., 2016; Von Ellenrieder et al., 2016; Hedrich et al., 2017; Fan et al., 2021; Guo et al., 2022) have shown that MEG can detect epileptic spikes and HFOs effectively. In the presurgical diagnosis process, it is critical to accurately detect the HFOs in MEG signals for improving the post-surgical outcomes of patients with epilepsy. Visual reviews of HFOs in MEG signals by human experts play an important role in current clinical practices. However, visual identification of HFOs is usually

subjective, time-consuming, and error prone due to the large volume of MEG signal data (Zelmann et al., 2012; Roehri et al., 2017; Fujiwara et al., 2020). Consequently, a number of automatic approaches (Gardner et al., 2007; Zelmann et al., 2010; Jacobs et al., 2012; Burnos et al., 2014) has been proposed to enable HFO detection so as to assist human experts for the visual review of iEEG and MEG signals. During the detection tasks, a universally two-step framework is applied by most of these methods: (1) The whole recording data is divided into a large number of signal segments. (2) The HFO detectors extract certain signal features for decision making. The handcrafted features that are manually designed based on observation or statistical analysis play as the solution for the feature of HFO signals. For example, Van Klink et al. (2017) proposed an automatic HFO detection and visualization approach in MEG. Similarly, in another work (Burnos et al., 2014), handcrafted features (e.g., high frequency peak and low frequency peak) were proposed to automatically distinguish HFOs in EEG signals. In these works, a cutoff for handcrafted features is often required to recognize an HFO signal segment. It is clear that these approaches based on handcrafted features require to be adjusted or re-optimized when the detectors are applied to similar neuroimaging data from different populations. This circumstance hinders the generalizability of HFOs in unseen conditions. Recently, machine learning provides a possible opportunity for improving the performance of HFO detections and reducing human interference. Traditional machine learning algorithm (Elahian et al., 2017), such as logistic regression, has been used for the identification of the epileptogenic zones. More recently, a deep learning approach SMO detector (Guo et al., 2018) was proposed. Such deep learning based HFO detector requires minimal human interference by using a golden standard dataset to train the detector.

Objective and automatic detection of HFOs in MEG signals with advanced deep learning algorithms may serve as a promising clinical decision support system to assist human experts for the visual review of MEG signals (Guo et al., 2018; Kong et al., 2019). In addition to the correct detection of HFOs in the MEG signals, recommending the possible results to clinicians is also crucial for an accurate and timely clinical evaluation. The recommended HFO list may not only serve as evidence for the particular patients but also serve as clinical diagnosis cases for future data retrieval purpose (e.g., teaching and research). In this study, our overall goal is set to develop a deep learning model to detect HFO signals with high confidence among a large amount of MEG data and recommend these findings to clinicians as a clinical decision support system. To that effect, we propose a multi-head self-attention deep neural network as the HFO detector. Compared to the existing algorithms (e.g., SMO detector Guo et al., 2018), we introduce the popular multi-head self-attention mechanism in this paper to enable an HFO detector to jointly pay attention to important

information from various representation subspaces at multiple positions. Instead of computing the attention once, this multi-head self-attention strategy is able to compute the importance of each feature multiple times in parallel. Our hypothesis is that an HFO detector with multi-head self-attention mechanism is able to outperform the existing detector based on deep neural networks. Our newly developed HFO detector enables clinicians to objectively and automatically observe and localize HFOs for the preparation of epilepsy surgery without human designed signal features. According to the output probability values of our proposed detector, the MEG signals can be sorted in descending order. In order to accurately and timely understand the patient's condition, we can recommend N signals with the highest HFO signal probability value (e.g., top-10) to the clinician and assist in developing a treatment plan. This process is also known as a top- N recommendation task (Zhao et al., 2013, 2014; Wang et al., 2017; Zhang et al., 2017).

The following sections of this paper are organized as follows: First, we describe the patients and their associated MEG data in this work and the detailed MSADR framework in the Section 2. Second, experiment setups, such as model evaluation, peer machine learning models, and developmental environment are described. Third, we present the performance of detection and recommendation of MSADR for HFO signals. Then, ablation studies are also conducted to test our MSADR approach. Fourth, the discoveries and limitation of this work are discussed. Finally, we conclude the paper by summarizing the contributions and future directions.

2. Materials and methods

2.1. MEG data

2.1.1. Data acquisition

In this retrospective study, we obtained interictal MEG data from 20 clinical patients with epilepsy consisting of 10 females and 10 males (age: 6–60 years, mean age: 32 years), who were affected by focal seizures arising from one part of the brain. The Institutional Review Board was approved, and written informed consents were obtained from all subjects.

Full details of MEG data acquisition can be found in our prior study (Guo et al., 2018). Briefly, MEG recordings were performed using a 306-channel, whole-head MEG system (VectorView, Elekta Neuromag, Helsinki, Finland) in a magnetically shielded room. Sleep deprivation and reduction of anti-epileptic drugs were used to increase the chance for capture HFOs during MEG recordings, as one part of the pre-surgical evaluation. An approximate 1 h of MEG data was recorded for all patients. The sampling rate of MEG data was 2,400 Hz. The noise floor in our MEG systems was calculated with MEG data acquired without a subject (empty room). The noise floor was used to identify MEG system noise. The noise level was

about 3–5 fT/Hz. The empty room measurements were also used to compute the noise covariance matrix for localizing epileptic activities (i.e., HFOs). A three-dimensional coordinate frame relative to the subject's head was derived from these positions. The system allowed head localization to an accuracy of 1 mm. The changes in head location before and after acquisition were required to be less than 5 mm for the study to be accepted. To identify the system and environmental noise, we routinely recorded one background MEG dataset without patients just before the experiment.

2.1.2. Segment

A public available software MEG Processor (Xiang et al., 2009) was used to correct and label the MEG data. In the current work, the MEG data were segmented into about 11,016,000 signal segments¹ with 2 s window size without overlap. Each signal segment was a MEG signal intensity vector with 4,800 data points in the time domain. These segments were first filtered automatically, which segments with a goodness-of-fit value of < 85% or confidence volume of > 3mm³ were dropped. Second, two band-pass filter, an 80–250 Hz one for ripples and a 250–500 Hz one for fast ripples, were introduced to filter high frequency MEG data into candidate segment set while the low frequency ones were automatically dropped. Both physiologic and pathologic high frequency neuromagnetic signals were included in the candidate segment set. The physiologic HFOs were manually rejected, and then the pathologic ones were selected by comparing MEG ripples and iEEG recordings at source levels (Wu et al., 2014) by two human experts. A total of 660 HFO signal segments selected by human experts, together with 660 normal control (NC) signal segments randomly selected from the rest segments, were compiled into our gold standard dataset. Figure 1 shows examples of MEG data, HFO, and NC segments.

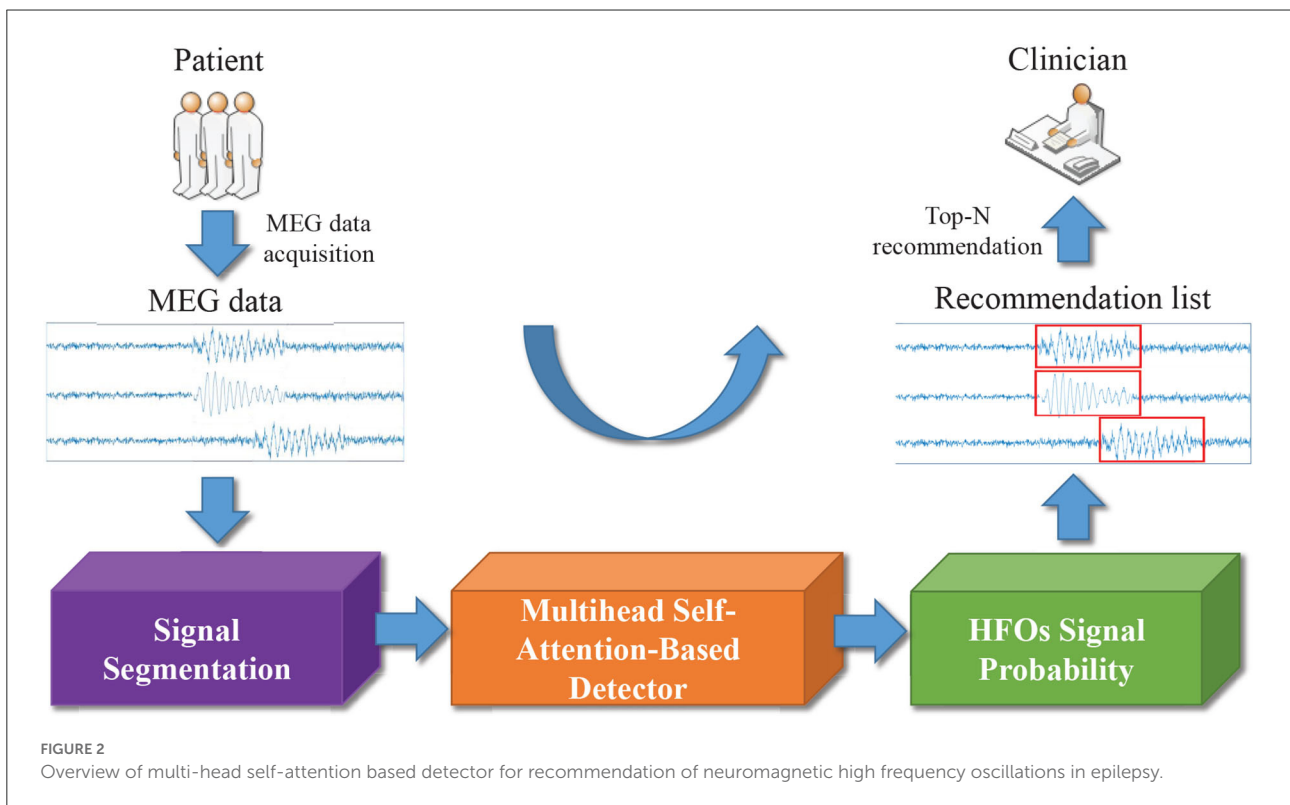
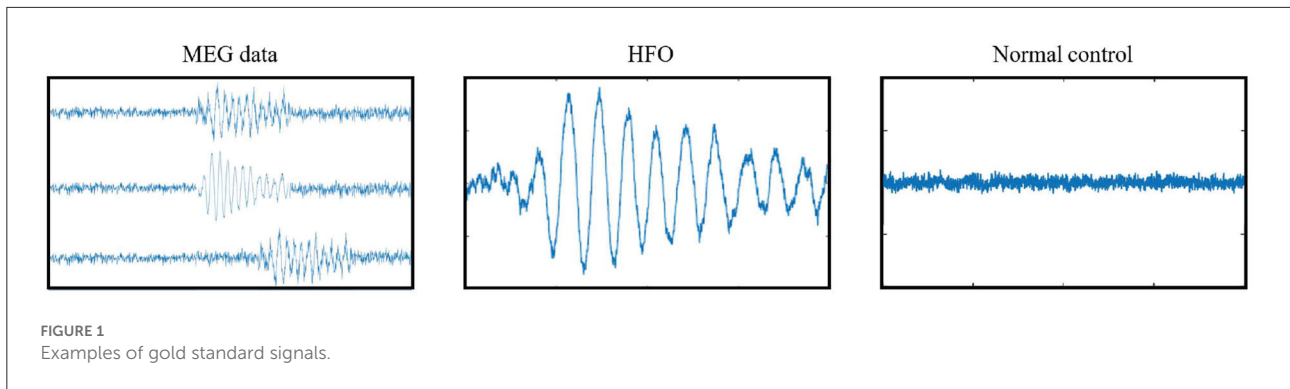
2.2. Method

2.2.1. Overview framework for HFO detection and recommendation

In Figure 2, we display the overview of our multi-head self-attention based detector for recommendation (MSADR), which consists of MEG data acquisition, signal segmentation (purple box), multi-head self-attention based detector (orange box), HFO signal probability (green box), and detection and recommendation list.

The structure of multi-head self-attention based detector (MSAD) is given in Figure 3. It consists of layers of dense,

¹ The number of segments is calculated as $20(\text{patients}) \times 306(\text{channels}) \times 60(\text{min}) \times 60(\text{s}) \div 2(\text{s}/\text{window})$.



normalization, multi-head self-attention, self-attention. The various components are described in the following sections.

2.2.2. Dropout, dense, and normalization

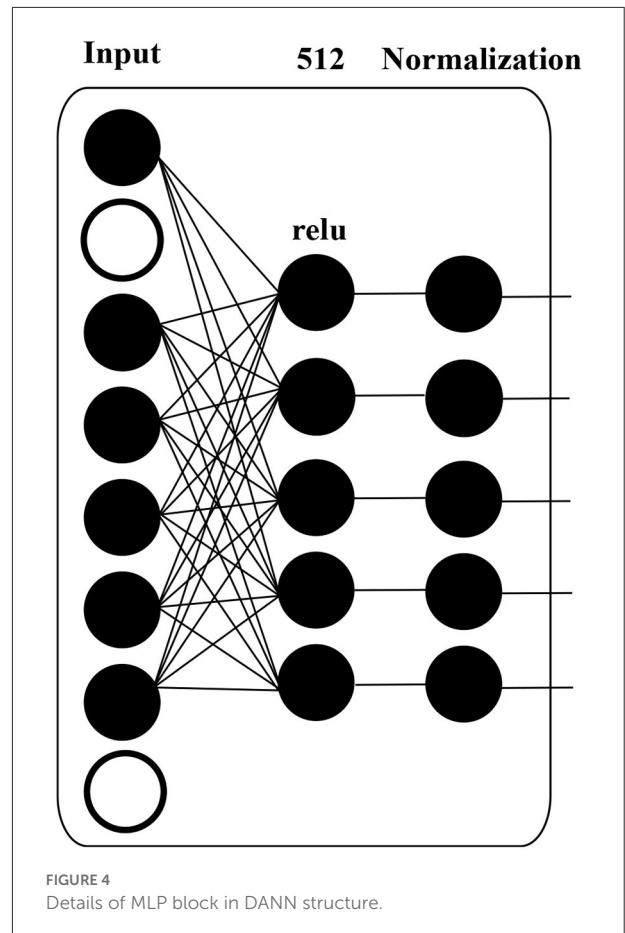
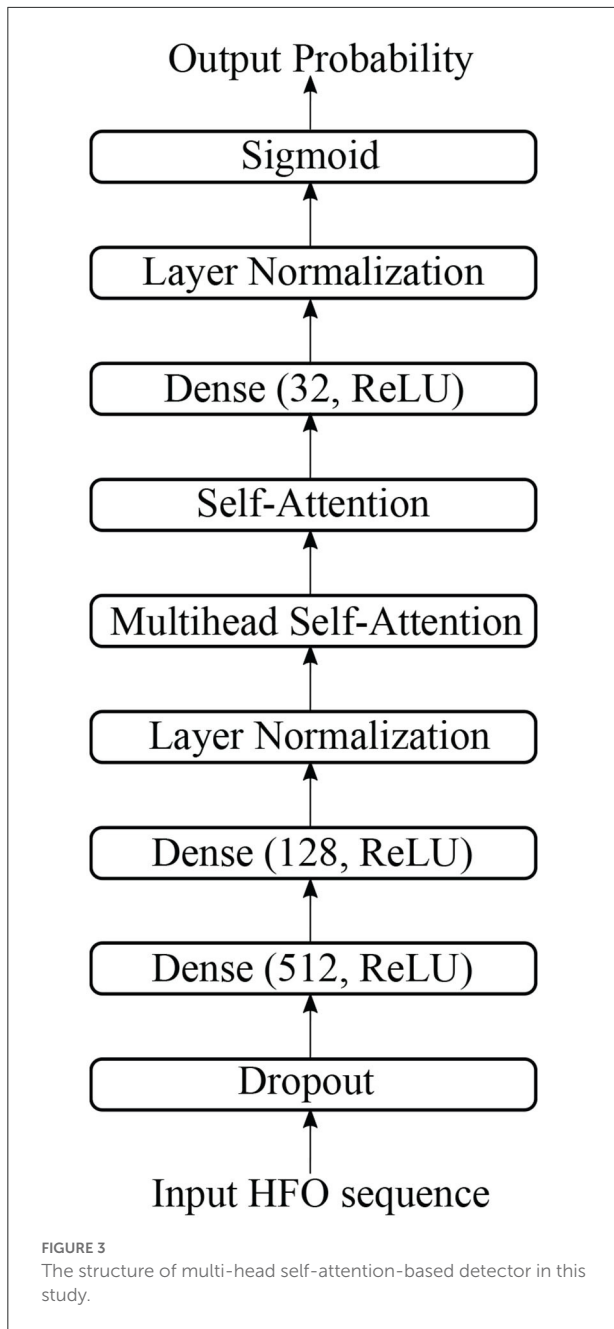
There is one dropout layer, three dense layers, and two normalization layers in our proposed model. Figure 4 shows the computation details of these layers.

A dropout layer, which prevents over-fitting during model training, is applied to input data, i.e., HFO sequence. The white circle in Figure 4 indicates dropped units according to dropout probability. The dropout layer is followed by a dense layer, whose hidden units are 512, and the activation function is “relu”, to reduce the dimension of the previous layer. The normalization

layer (Ioffe and Szegedy, 2015) is used to accelerate deep network training by reducing internal covariate shift.

2.2.3. Self-attention

The attention is proposed to compute an alignment score between elements from two sources (Shen et al., 2018). In particular, given a sequence of HFOs, $\mathbf{x} = [x_1, x_2, \dots, x_n]$ and a representation of a query $q \in \mathbb{R}^d$, the attention computes the alignment score between q and each element x_i using a compatibility function $f(x_i, q)$. A softmax function then transforms the alignment scores $[f(x_i, q)]_{i=1}^n$ to a probability distribution $p(z|\mathbf{x}, q)$, where z represents the importance degree to q . That is, a large $p(z = i|\mathbf{x}, q)$ means that x_i contributes



Additive attention (Bahdanau et al., 2015; Shang et al., 2015) is a commonly-used attention mechanism where the compatibility function $f(\cdot)$ is parameterized by a MLP, i.e.:

$$f(x_i, q) = w^T \sigma(W^{(1)}x_i + W^{(2)}q), \tag{4}$$

where $W^{(1)} \in \mathbb{R}^{d \times d}$, $W^{(2)} \in \mathbb{R}^{d \times d}$, $w \in \mathbb{R}^d$ are learnable parameters, d is the number of columns of x_i , and $\sigma(\cdot)$ is an activation function. Compared with multiplicative attention (Rush et al., 2015; Sukhbaatar et al., 2015) using cosine similarity or inner product as the compatibility function for $f(x_i, q)$, i.e.:

$$f(x_i, q) = \langle W^{(1)}x_i, W^{(2)}q \rangle, \tag{5}$$

Though additive attention is expensive in time cost and memory consumption, it achieves better empirical performance for downstream tasks.

Self-attention (Liu et al., 2016; Lin et al., 2017; Peng et al., 2021) explores the importance of each feature to the entire HFOs given a specific task. In particular, q is removed from the common compatibility function which is formally written as the following equation:

$$f(x_i) = w^T \sigma(W^{(1)}x_i), \tag{6}$$

important information to q . This attention process can be formalized as follows:

$$\alpha = [f(x_i, q)]_{i=1}^n, \tag{1}$$

$$p(z = i|x, q) = \text{softmax}(\alpha). \tag{2}$$

The output *Attention* is the weighted element according to its importance, i.e.,

$$\text{Attention}(q, \mathbf{x}) = p(z = i|x, q)\mathbf{x}. \tag{3}$$

$$\alpha = [f(x_i)]_{i=1}^n, \tag{7}$$

$$p(z = i|\mathbf{x}) = \text{softmax}(\alpha). \tag{8}$$

The output *Attention* is the weighted element according to its importance, i.e.,

$$\text{Attention}(\mathbf{x}) = p(z = i|\mathbf{x})\mathbf{x}. \tag{9}$$

2.2.4. Multi-head self-attention

Multi-head self-attention allows the model to jointly attend to information from different representation subspaces at different positions. We use the multi-head version with k heads, as introduced in Vaswani et al. (2017),

$$\text{MultiHead}(\mathbf{x}) = \text{Concat}(\text{head}_1, \dots, \text{head}_k)W^{(O)}, \tag{10}$$

$$\text{where } \text{head}_i = \text{Attention}(\mathbf{x}W^{(x)}), \tag{11}$$

where projections using learned parameter matrices $W^{(x)} \in \mathbb{R}^{d \times d/k}$, and $W^{(O)} \in \mathbb{R}^{d \times d}$.

2.2.5. Loss function

A standard cross-entropy loss is used as the training objective of MSADR, defined as

$$\mathcal{L} = -y \log(p) - (1 - y) \log(1 - p), \tag{12}$$

where y is the target label (0 or 1) and p is the predicted probability between 0 and 1 given an HFO sequence.

3. Experiment setup

We evaluate the proposed model on two tasks including the classification of whole patients and recommendation for each individual patient.

3.1. Model evaluation

We conduct a comprehensive evaluation in this study by employing the proposed MSADR on the HFO dataset to classify HFO data and recommend detected HFO sequences to medical experts. We employ the evaluation strategy of leave-one-out cross-validation in our experiments. In the 10-fold leave-one-out cross-validation, the HFO dataset is separated to two parts. One consists of 90% of the whole as training data while the rest part is regarded as test data.

For the detection task, we put all segments from 20 patients together to separate the training data and test data. We first calculate true positive (TP), false positive (FP), true negative

(TN), and false negative (FN) by comparing the predicted labels and gold-standard labels. Then, we calculate accuracy, recall, precision, and F-score by

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN},$$

$$\text{Recall} = \frac{TP}{TP + FN},$$

$$\text{Precision} = \frac{TP}{TP + FP},$$

$$F\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

For the recommendation task, the dataset is separated according to patients, in each split, segments from 18 patients are selected into the training set while segments from the rest 2 patients as test data. We use top- N precision (P@ N) (Choi et al., 2016) to evaluate the ability of the algorithm to recommend detected HFOs for individual patients in the test set, defined as follows:

$$P@N = \frac{TP@N}{N},$$

where ‘‘TP@ N ’’ in the formula stands for TP HFOs in top- N recommendation task. Top- N precision mimics the behavior of doctors conducting differential diagnoses, where doctors list most probable diagnoses and treat patients accordingly to identify the patient status. Therefore, a machine with a high top- N precision translates to a doctor with an effective diagnostic skill. This makes top- N precision an attractive performance metric for our problem (Choi et al., 2016).

3.2. Peer machine learning models

To compare our proposed model MSADR with existing machine learning models, we also implemented random forest (RF) (Breiman, 2001; Nissen et al., 2018), support vector machine (SVM) models (Ak et al., 1999; Zhang et al., 2020), and SMO detector (Guo et al., 2018).

- **Random Forest (RF):** RF is a classic ensemble learning methods by learning multiple decision trees and employing averaging to improve classification performance and control over-fitting. The number of trees in the forest was optimized from empirical values [20, 40, 60, 80, and 100]. We set maximal depth of the tree as 10.
- **Support Vector Machine (SVM):** A SVM model is developed to perform classification by using vectorized FC features. We apply a linear kernel and search for the margin penalty with empirical values [0.2, 0.4, 0.6, 0.8, and 1.0].
- **SMO detector (Guo et al., 2018):** In terms of the existing deep learning model, we compared our model with SMO

detector, a DNN model developed previously for HFO detection. Briefly, we implemented the SMO detector model as a 7-layer DNN, with input number of HFO sequence in the input layer, followed by dropout, dense (512, ReLU), dense (128, ReLU), normalization, dense (32, ReLU), normalization, and the sigmoid layer to generate one output unit. A cross entropy loss function is applied to supervise the network learning adopted. The learning rate is set as 0.0001. A total of 10 epochs are applied to ensure the convergence of the model.

3.3. Developmental environment

The proposed DANN and peer machine learning models are all implemented in Python 3.7 environment. To build the deep learning related models, we apply TensorFlow (2.0.0-rc1) backend. For the traditional models, we adopt the models from Sklearn 0.20.2.

All the experiments are conducted on a workstation with 10 cores of Intel Core i9 CPU and 64GB RAM. Due to the high computation cost of deep learning algorithm, we use one GPU (Nvidia TITAN Xp, 12GB RAM) to accelerate the training speed of the models.

4. Results

We evaluate the performance of detection and recommendation for each set of experiments. There are two sets of experiments to be conducted, which consist of overall performance compared with baseline models and effectiveness of varying head number of multi-head self-attention.

4.1. Overall performance comparison

4.1.1. Detection

We first compare the HFO detection performance of the proposed MSADR model and multiple peer machine learning models, including RF, SVM, and SMO. The results are derived on a leave-one-out cross-validation experiment by using the entire dataset. As shown in Table 1, our proposed MSADR takes the lead place (the bold value) in all metrics of HFO detection accuracy (0.886), recall (0.840), and F-score (0.859) among compared machine learning models, while the RF model returns the lowest detection performance on recall and F-score, and SVM on accuracy. Our model outperforms the SVM model by 0.126 on accuracy, the RF model by 0.263 on recall, and 0.142 on F-score.

4.1.2. Recommendation

We then compare the HFO recommendation performance of the proposed MSADR model and baseline models including RF, SVM, and SMO. The experiment setting is almost the same as the detection task, except for an additional recommendation module to generate a ranking list. As shown in Table 2, our proposed MSADR obtains the best HFO recommendation performance on P@1 (0.967), P@3 (0.858), and P@5 (0.879) among the compared machine learning models, whereas the RF model returns the lowest recommendation performance on P@3, P@5, and SVM on P@1. Our model increases the performance of SVM by 0.1 on P@1, the RF model by 0.162 and 0.215 on P@3 and P@5, respectively.

4.1.3. Computational costs

The computational costs of the proposed MSADR model and baseline models are provided in Table 3. It is noticed that our proposed MSADR takes a longer time for model training and inference than baseline models. The main reason is the complexity of the deep learning models (MSADR and SMO) and the huge number of the parameters. The MSADR has 411,325 parameters to be trained, which obtains a stronger learning capacity to get the best model performance. MSADR uses a learned model to conduct the detection and recommendation task, which theoretically take more time.

The results in Tables 1, 2 also show a trend that deep learning models (MSADR and SMO) achieve improved performance compared to the traditional model, such as SVM and RF, demonstrating the superior capability of the deep learning

TABLE 1 Detection comparison of random forest (RF), support vector machine (SVM), SMO, and multi-head self-attention-based detector for recommendation (MSADR) trained using leave-one-out cross-validation on the entire dataset.

Method	Accuracy	Recall	Precision	F-Score
RF	0.779	0.577	0.951	0.717
SVM	0.760	0.743	0.764	0.753
SMO detector (Guo et al., 2018)	0.845	0.732	0.951	0.826
MSADR	0.886	0.840	0.881	0.859

The bold values represent the lead place of the corresponding metric.

TABLE 2 Recommendation comparison of RF, SVM, SMO, and MSADR trained using leave-one-out cross-validation on the entire dataset.

Method	P@1	P@3	P@5
RF	0.893	0.696	0.664
SVM	0.867	0.800	0.760
SMO detector (Guo et al., 2018)	0.893	0.811	0.793
MSADR	0.967	0.858	0.879

The bold values represent the lead place of the corresponding metric.

TABLE 3 Computational cost comparison of RF, SVM, SMO, and MSADR.

Method	Training time (ms)	Inference time (ms)	No. of parameters
RF	284	6	-
SVM	83	5	2,400
SMO detector (Guo et al., 2018)	2,506	140	318,449
MSADR	12,317	525	411,325

model on complex data patterns, such as HFO. In addition, the inference time of MSADR is about 0.5 s (525 ms). This is an acceptable time cost while it can bring about 16.6% accuracy improvement on detection and 11.5% P@1 improvement on recommendation toward the fastest model (SVM). Another trend can be observed that the P@1 returns the best recommendation score a cross all models, which demonstrates that the machine learning model is a promising alternative approach to assist clinicians to make decisions.

4.2. Effectiveness of varying head number of multi-head self-attention

The effectiveness of our MSADR is further tested by varying head number ($k = [2, 4, 8, 16]$) of multi-head self-attention on two tasks of detection and recommendation. The results in this set of experiments are calculated based on leave-one-out cross-validation by using the entire dataset.

4.2.1. Detection

The HFO detection performance of the proposed MSADR model has been evaluated by varying head number of multi-head self-attention. Figure 5A displays plots of the accuracy, recall, precision, and F-score of the proposed MSADR over different strategies of varying head number.

It is apparent that the proposed MSADR achieves the state-of-the-art in terms of accuracy, recall, and F-score when the head number is set as 8, while the performance on precision is the worst. Performance accuracy, recall, and F-score decrease as the head number increase due to the over-fitting of self-attention, whereas performance on precision slightly increases. Overall, the proposed MSADR has the best three indicators out of four when the head number of multi-head self-attention is set to 8.

4.2.2. Recommendation

The effectiveness of the proposed MSADR model over HFO recommendation task has been compared in terms of P@N by varying head number of multihead self-attention. Figure 5B displays plots of the P@1, P@3, and P@5 of MSADR

over different strategies of varying head number. As shown in the figure, our proposed MSADR achieves the highest HFO recommendation performance on P@1 and P@5, while the performance of the two indicators decreases as the head number increases due to the over-fitting of multi-head self-attention. Overall, the proposed MSADR has the best two indicators out of three when the head number of multi-head self-attention is set to 8.

4.3. Ablation study

A detailed ablation study is performed to examine the contributions of the model's components to the tasks of detection and recommendation. There are four configurations of replaceable components in this model. The two components are (1) multi-head self-attention layer and (2) the self-attention layer. The four configurations based on MSADR are

- **raw (DNN):** (1) and (2) are removed from MSADR, which becomes a pure DNN, one of our peer baseline models (SMO detector);
- **Attn_1:** (2) is removed and (1) is remained in MSADR;
- **Attn_2:** (1) is removed and (2) is remained in MSADR;
- **MSADR:** our proposed model.

All models are trained with 10 epochs and a batch size of 32. The head number of multi-head self-attention is empirically set to 8.

4.3.1. Detection

From Table 4, we find that the MSADR model obtains the best performance on detection task compared to the ablated models, except for the performance of Attn_1 on precision. Moreover, we note that Attn_1 and Attn_2 outperform raw, which gives us the confidence to apply self-attention to learn the relationship between HFO signals. It is clear that the single self-attention model provides comparable information to the performance of the Attn_1 and Attn_2 model. In particular, MSADR outperforms the best ablated model for Accuracy by 1.1%, for Recall by 2.1%, and for F-score by 1.0%.

4.3.2. Recommendation

Table 5 shows the recommendation performance for the ablated models and our proposed model. As can be seen from the table, the proposed model achieves the best performance compared to the ablated models on the recommendation task. We observe that Attn_1 and Attn_2 outperform raw on P@1 and P@3, which again demonstrates that self-attention is a vital component to learn the relationship

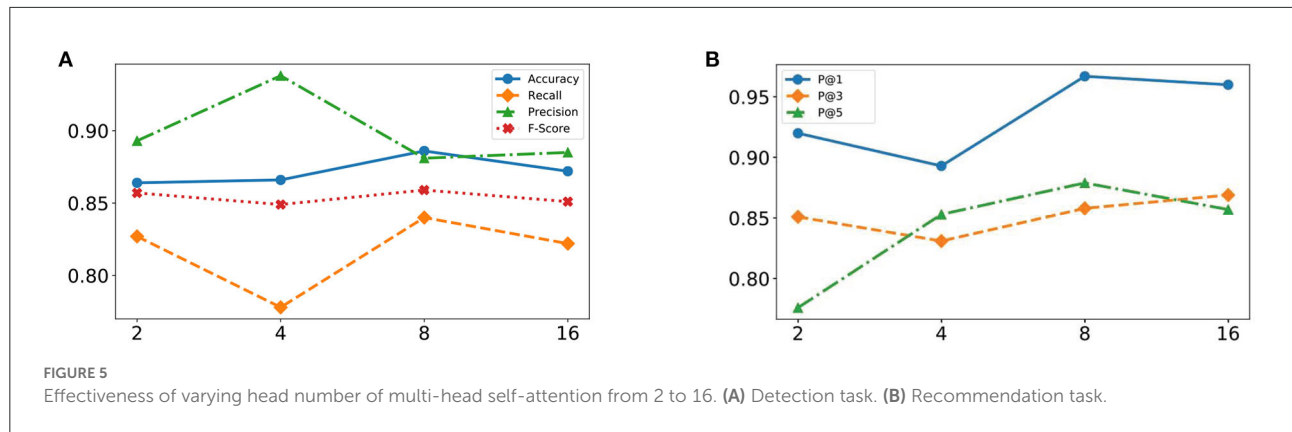


FIGURE 5 Effectiveness of varying head number of multi-head self-attention from 2 to 16. (A) Detection task. (B) Recommendation task.

TABLE 4 Detection comparison of ablated models trained using leave-one-out cross-validation on the entire dataset.

Method	Accuracy	Recall	Precision	F-Score
raw	0.845	0.732	0.951	0.826
Attn_1	0.845	0.755	0.975	0.849
Attn_2	0.875	0.819	0.928	0.847
MSADR	0.886	0.840	0.881	0.859

The bold values represent the lead place of the corresponding metric.

TABLE 5 Recommendation comparison of ablated models trained using leave-one-out cross-validation on the entire dataset.

Method	P@1	P@3	P@5
raw	0.893	0.811	0.793
Attn_1	0.913	0.844	0.759
Attn_2	0.920	0.847	0.873
MSADR	0.967	0.858	0.879

The bold values represent the lead place of the corresponding metric.

between HFO signals. MSADR outperforms the best ablated model by 4.7%, 1.1%, and 0.6% on P@1, P@3, and P@5, respectively.

5. Discussion

Since first discovered in the 1990s, HFOs have been considered a promising biomarker to locating the seizure onset zone and improving postsurgical outcomes in patients with epilepsy (Huang and White, 1989; Fan et al., 2021). Noninvasive brain recording technologies (i.e., scalp EEG and MEG) were a milestone in human HFO research and have provided the possibility to investigate this brain activity in a wider range (Papadelis et al., 2009, 2016; Van Klink et al., 2016; Von Ellenrieder et al., 2016; Hedrich et al., 2017). Due to excellent temporal resolution and acceptable spatial resolution, MEG is able to effectively record HFOs and localize

epileptic activities for epilepsy surgery (Fan et al., 2021). After noninvasive recording, the detection of HFOs is the next crucial task for onset zone detection. Although visual identification is still considered to be the gold standard for HFO detection, it still faces the problem of highly time-consuming and subjective (Frauscher et al., 2017).

This study mainly focuses on the automatic detection and recommendation of HFOs from interictal MEG data. The MEG data of clinical epileptic patients were recorded with a multi-channel whole-head MEG system (Xiang et al., 2010; Guo et al., 2018), and then segmented into signal segments with 2 s window size without overlap. The labeled HFO segments by human experts and randomly selected NC segments from the complementary set of labeled HFO set was compiled into our gold standard dataset. With the gold standard data, we trained the proposed MSADR algorithm for the detection and recommendation model of HFOs. For a new patient, the trained model can detect HFOs from the segmented MEG data and recommend HFO signals to clinicians, alleviating the burden on reviewing the large amount of MEG data. The effectiveness of our proposed detection and recommendation approaches were demonstrated by the cross-validation experimental results. The proposed MSADR can improve the detection accuracy by at least 13.7% and the top-1 recommendation precision by 8.2% compared with the traditional machine learning methods (RF and SVM) while improving the detection accuracy by 4.8% and the top-1 recommendation precision by 8.2% compared with another deep learning method (SMO detector). The computational costs are important, especially in real world applications. Though MSADR is a time-consuming method, the acceptable inference time (0.5 s) can guarantee the user experience. In addition, sliding window with overlap can be used for segmenting, so as to improve the possibility of HFO locating in the center of segments in real world applications.

There are some limitations to this study. First, the experiment was built on a small dataset with 20 patients. A larger data set is required to further validate the effectiveness and efficiency of MSADR. Second, this work only focused on

the detection and recommendation of HFOs from interictal MEG data. Performance of the MSADR approach on other neuromagnetic data (i.e., ictal MEG, iEEG, and EEG) remains unclear. We will test our method in future work. Third, the MEG segments from different patients or channels are treated equally and independently in this paper. However, there are complex timing and co-occurrence relationships among segments. Mining and utilizing these relationships may improve the effectiveness of HFO detection and recommendation. Finally, since the current approach requires signal segmentation of MEG data, it is only able to differentiate HFOs and NC segments with a pre-defined fixed signal length. It cannot directly detect HFOs in an automatic way on the raw MEG data (e.g., start and end positions).

Due to the high cost and lack of automatic detection technology with broad applicability, traditional MEG has limited availability (Guo et al., 2018; Kong et al., 2019). However, the technological innovations in MEG have been progressing. New MEG systems with optically pumped magnetometers do not require cooling with liquid helium and can be worn more conveniently (Boto et al., 2019). This may reduce the cost of MEG data recording and expand the scope of application. The research and application of automatic or semi-automatic HFO detection methods with broad applicability will make more efficient use of MEG data (Guo et al., 2018), and the integration into clinical review software can effectively enhance clinical value, including preoperative localization of epileptogenic regions, the assessment of disease severity, predicting seizures, monitoring treatment, evaluating treatment effects, and assessing epileptic susceptibility after brain injury (Fan et al., 2021).

6. Conclusion

In this study, we develop an MSADR detector for the detection and recommendation of HFO signals by using the multi-head self-attention mechanism. By comparing our model with traditional machine learning models (RF and SVM) and deep learning model (SMO detector), the proposed MSADR detector is proved to reach state-of-the-art performance in both detection and recommendation tasks. The robustness of our detector is also extensively assessed with multiple ablation tests. The MSADR is supposed to detect HFOs from a large amount of segmented MEG data and recommend HFO signals to help clinicians locate epileptogenic regions and assist in treatment. Our future directions may focus on extending our model to capture the timing and co-occurrence relationships among segments to further improve the effectiveness of HFO detection and recommendation.

Data availability statement

The datasets presented in this article are not readily available because of a confidentiality agreement that prevents them from being disclosed to the public. Requests to access the datasets should be directed to JG, guojy@xmu.edu.cn.

Ethics statement

The studies involving human participants were reviewed and approved by Nanjing Brain Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

XZ, XP, KN, JG, and YP contributed to the conception and design of the study. XP and FY developed the DL algorithms. TW made a contribution to data collection. XZ, DC, QZ, MO, and JG performed the statistical analysis. XZ and KN wrote the initial draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This work was supported by the National Key Research and Development Program of China (Nos. 2020YFD1100605 and 2019YFD1101100), Beijing Information Science and Technology University Qin Xin Talents Cultivation Program (No. QXTCP C202112), Research Level Improvement Project (No. 2020KYNH214), Beijing Educational Science Planning Project of China (No. CHCA2020102), National Institutes of Health (Nos. R01-EB029944 and R01-EB030582), National Natural Science Foundation of China (Nos. 82172022 and 62006100), 2025 Key Technology Innovation Program of Ningbo City (No. 2019B10082) Fundamental Research Funds for the Chinese Central Universities (0070ZK1096 to JG), and Ningbo Public Welfare Technology Plan Project (No. 2019C50081).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AM declared a shared affiliation with the author TW to the handling editor at the time of review.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Ak, J., Suykens, J., and Vandewalle (1999). Least squares support vector machine classifiers. *Neural Process. Lett.* 9, 293–300.
- Bahdanau, D., Cho, K. H., and Bengio, Y. (2015). “Neural machine translation by jointly learning to align and translate,” in *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015* (San Diego, CA).
- Boto, E., Seedat, Z. A., Holmes, N., Leggett, J., Hill, R. M., Roberts, G., et al. (2019). Wearable neuroimaging: combining and contrasting magnetoencephalography and electroencephalography. *Neuroimage* 201, 116099–116099. doi: 10.1016/j.neuroimage.2019.116099
- Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Burnos, S., Hilfiker, P., Sürücü, O., Scholkmann, F., Kraysenbühl, N., Grunwald, T., et al. (2014). Human intracranial high frequency oscillations (hfos) detected by automatic time-frequency analysis. *PLoS ONE* 9, e94381. doi: 10.1371/journal.pone.0094381
- Choi, E., Bahadori, M. T., Schuetz, A., Stewart, W. F., and Sun, J. (2016). “Doctor ai: predicting clinical events via recurrent neural networks,” in *Machine Learning for Healthcare Conference* (Los Angeles, CA), 301–318.
- Durnford, A. J., Rodgers, W., Kirkham, F. J., Mullee, M. A., Whitney, A., Prett, M., et al. (2011). Very good inter-rater reliability of engel and ilae epilepsy surgery outcome classifications in a series of 76 patients. *Seizure* 20, 809–812. doi: 10.1016/j.seizure.2011.08.004
- Elahian, B., Yeasin, M., Mudigoudar, B., Wheless, J. W., and Babajani-Feremi, A. (2017). Identifying seizure onset zone from electrocorticographic recordings: a machine learning approach based on phase locking value. *Seizure* 51, 35–42. doi: 10.1016/j.seizure.2017.07.010
- Fan, Y., Dong, L., Liu, X., Wang, H., and Liu, Y. (2021). Recent advances in the noninvasive detection of high-frequency oscillations in the human brain. *Rev. Neurosci.* 32, 305–321. doi: 10.1515/revneuro-2020-0073
- Frauscher, B., Bartolomei, F., Kobayashi, K., Cimbalk, J., van 't Klooster, M. A., Rampp, S., et al. (2017). High-frequency oscillations: the state of clinical research. *Epilepsia* 58, 1316–1329. doi: 10.1111/epi.13829
- Fujiwara, K., Huang, Y., Hori, K., Nishioji, K., Kobayashi, M., Kamaguchi, M., et al. (2020). Over- and under-sampling approach for extremely imbalanced and small minority data problem in health record analysis. *Front. Public Health* 8, 178. doi: 10.3389/fpubh.2020.00178
- Gardner, A. B., Worrell, G. A., Marsh, E., Dlugos, D., and Litt, B. (2007). Human and automated detection of high-frequency oscillations in clinical intracranial eeg recordings. *Clin. Neurophysiol.* 118, 1134–1143. doi: 10.1016/j.clinph.2006.12.019
- Guo, J., Xiao, N., Li, H., He, L., Li, Q., Wu, T., et al. (2022). Transformer-based high-frequency oscillation signal detection on magnetoencephalography from epileptic patients. *Front. Mol. Biosci.* 9, 822810–822810. doi: 10.3389/fmolb.2022.822810
- Guo, J., Yang, K., Liu, H., Yin, C., Xiang, J., Li, H., et al. (2018). A stacked sparse autoencoder-based detector for automatic identification of neuromagnetic high frequency oscillations in epilepsy. *IEEE Trans. Med. Imaging* 37, 2474–2482. doi: 10.1109/TMI.2018.2836965
- Hedrich, T., Pellegrino, G., Kobayashi, E., Lina, J.-M., and Grova, C. (2017). Comparison of the spatial resolution of source imaging techniques in high-density eeg and meg. *Neuroimage* 157, 531–544. doi: 10.1016/j.neuroimage.2017.06.022
- Hu, J., Wang, C. S., Wu, M., Du, Y. X., He, Y., and She, J. (2015). Removal of eeg and emg artifacts from eeg using combination of functional link neural network and adaptive neural fuzzy inference system. *Neurocomputing* 151, 278–287. doi: 10.1016/j.neucom.2014.09.040
- Huang, C., and White, L. (1989). High-frequency components in epileptiform EEG. *J. Neurosci. Methods* 30, 197–201. doi: 10.1016/0165-0270(89)90130-1
- Ioffe, S., and Szegedy, C. (2015). “Batch normalization: accelerating deep network training by reducing internal covariate shift,” in *Proceedings of the 32nd International Conference on Machine Learning* (Lille), 448–456.
- Jacobs, J., Staba, R., Asano, E., Otsubo, H., Wu, J., Zijlmans, M., et al. (2012). High-frequency oscillations (hfos) in clinical epilepsy. *Progr. Neurobiol.* 98, 302–315. doi: 10.1016/j.pneurobio.2012.03.001
- Kong, Y., Gao, J., Xu, Y., Pan, Y., Wang, J., and Liu, J. (2019). Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier. *Neurocomputing* 324, 63–68. doi: 10.1016/j.neucom.2018.04.080
- Leung, H., Poon, W. S., Kwan, P. K., Lui, C. H., Poon, T. L., Chan, E. L., et al. (2018). Ictal intracranial electroencephalography using wavelet analysis of high-frequency oscillations in chinese patients with refractory epilepsy. *Hong Kong Med. J.* 24, 21–23.
- Lin, Z., Feng, M., dos Santos, C. N., Yu, M., Xiang, B., Zhou, B., et al. (2017). A structured self-attentive sentence embedding. *arXiv:1703.03130*. doi: 10.48550/arXiv.1703.03130
- Liu, Y., Sun, C., Lin, L., and Wang, X. (2016). Learning natural language inference using bidirectional lstm model and inner-attention. *arXiv:1605.09090*. doi: 10.48550/arXiv.1605.09090
- Miao, A., Xiang, J., Tang, L., Ge, H., Liu, H., Wu, T., et al. (2014). Using ictal high-frequency oscillations (80–500 hz) to localize seizure onset zones in childhood absence epilepsy: a meg study. *Neurosci. Lett.* 566, 21–26. doi: 10.1016/j.neulet.2014.02.038
- Modur, P. N. (2014). High frequency oscillations and infraslow activity in epilepsy. *Ann. Indian Acad. Neurol.* 17(Suppl. 1), S99. doi: 10.4103/0972-2327.128674
- Nakasato, N., Levesque, M. F., Barth, D. S., Baumgartner, C., Rogers, R. L., and Sutherling, W. W. (1994). Comparisons of meg, eeg, and ecog source localization in neocortical partial epilepsy in humans. *Electroencephalogr., Clin. Neurophysiol.* 91, 171–178. doi: 10.1016/0013-4694(94)90067-1
- Nevalainen, P., von Ellenrieder, N., Klimes, P., Dubeau, F., Frauscher, B., and Gotman, J. (2020). Association of fast ripples on intracranial eeg and outcomes after epilepsy surgery. *Neurology* 95, 10468. doi: 10.1212/WNL.00000000000010468
- Nissen, I. A., Stam, C. J., Van, S., Viktor, W., Reijneveld, J. C., Baayen, J. C., et al. (2018). Localization of the epileptogenic zone using interictal meg and machine learning in a large cohort of drug-resistant epilepsy patients. *Front. Neurol.* 9, 647. doi: 10.3389/fneur.2018.00647
- Olan Çocuklarda, M. N. D. E., and Öncesi, C. (2015). Presurgical evaluation and epilepsy surgery in mri negative resistant epilepsy of childhood with good outcome. *Turk Neurosurg* 25, 905–913. doi: 10.5137/1019-5149.JTN.12093-14.0
- Oldham, M. S., Horn, P. S., Tsevat, J., and Standridge, S. (2015). Costs and clinical outcomes of epilepsy surgery in children with drug-resistant epilepsy. *Pediatr. Neurol.* 53, 216–220. doi: 10.1016/j.pediatrneurol.2015.05.009
- Ontario, H. Q. (2012). Epilepsy surgery: an evidence summary. *Ont. Health Technol. Assess. Ser.* 12, 1.
- Papadelis, C., Poghosyan, V., Fenwick, P. B., and Ioannides, A. A. (2009). Meg's ability to localise accurately weak transient neural sources. *Clin. Neurophysiol.* 120, 1958–1970. doi: 10.1016/j.clinph.2009.08.018
- Papadelis, C., Tamilya, E., Stufflebeam, S., Grant, P. E., Madsen, J. R., Pearl, P. L., et al. (2016). Interictal high frequency oscillations detected with simultaneous magnetoencephalography and electroencephalography as biomarker of pediatric epilepsy. *J. Vis. Exp.* 6, e54883. doi: 10.3791/54883
- Peng, X., Long, G., Shen, T., Wang, S., and Jiang, J. (2021). “Sequential diagnosis prediction with transformer and ontological representation,” in *2021 IEEE International Conference on Data Mining (ICDM)* (Auckland: IEEE), 489–498.
- Reinholdson, J., Olsson, I., Edelvik, A., Hallböök, T., Lundgren, J., Rydenhag, B., et al. (2015). Long-term follow-up after epilepsy surgery in infancy and early childhood—a prospective population based observational study. *Seizure* 30, 83–89. doi: 10.1016/j.seizure.2015.05.019

- Roehri, N., Pizzo, F., Bartolomei, F., Wendling, F., and Bénar, C.-G. (2017). What are the assets and weaknesses of hfo detectors? a benchmark framework based on realistic simulations. *PLoS ONE* 12, e0174702. doi: 10.1371/journal.pone.0174702
- Rush, A. M., Chopra, S., and Weston, J. (2015). A neural attention model for abstractive sentence summarization. *arXiv:1509.00685*. doi: 10.18653/v1/D15-1044
- Shang, L., Lu, Z., and Li, H. (2015). "Neural responding machine for short-text conversation," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (Beijing: Association for Computational Linguistics), 1577–1586.
- Shen, T., Zhou, T., Long, G., Jiang, J., Pan, S., and Zhang, C. (2018). "Disan: directional self-attention network for rnn/cnn-free language understanding," in *AAAI'18/IAAI'18/EAAI'18: Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence* (New Orleans, LA).
- Stigsdotter-Broman, L., Olsson, I., Flink, R., Rydenhag, B., and Malmgren, K. (2014). Long-term follow-up after callosotomy—a prospective, population based, observational study. *Epilepsia* 55, 316–321. doi: 10.1111/epi.12488
- Sukhbaatar, S., Weston, J., and Fergus, R. (2015). "End-to-end memory networks," in *NIPS* (Montreal, QC), 2440–2448.
- Tamilia, E., Madsen, J. R., Grant, P. E., Pearl, P. L., and Papadelis, C. (2017). Current and emerging potential of magnetoencephalography in the detection and localization of high-frequency oscillations in epilepsy. *Front. Neurol.* 8, 14. doi: 10.3389/fneur.2017.00014
- Van Klink, N., Hillebrand, A., and Zijlmans, M. (2016). Identification of epileptic high frequency oscillations in the time domain by using meg beamformer-based virtual sensors. *Clin. Neurophysiol.* 127, 197–208. doi: 10.1016/j.clinph.2015.06.008
- Van Klink, N., Van Rosmalen, F., Nenonen, J., Burnos, S., Helle, L., Taulu, S., et al. (2017). Automatic detection and visualisation of meg ripple oscillations in epilepsy. *Neuroimage Clin.* 15, 689–701. doi: 10.1016/j.nicl.2017.06.024
- Van Klink, N., Van't Klooster, M., Zelmann, R., Leijten, F., Ferrier, C., Braun, K., et al. (2014). High frequency oscillations in intra-operative electrocorticography before and after epilepsy surgery. *Clin. Neurophysiol.* 125, 2212–2219. doi: 10.1016/j.clinph.2014.03.004
- Van't Klooster, M. A., Leijten, F. S., Huiskamp, G., Ronner, H. E., Baayen, J. C., Van Rijen, P. C., et al. (2015). High frequency oscillations in the intra-operative ecog to guide epilepsy surgery ("the hfo trial"): study protocol for a randomized controlled trial. *Trials* 16, 422. doi: 10.1186/s13063-015-0932-6
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). "Attention is all you need," in *NeurIPS* (Long Beach, CA), 5998–6008.
- Verdinelli, C., Olsson, I., Edelvik, A., Hallböök, T., Rydenhag, B., and Malmgren, K. (2015). A long-term patient perspective after hemispherotomy—a population based study. *Seizure* 30, 76–82. doi: 10.1016/j.seizure.2015.05.016
- Von Ellenrieder, N., Pellegrino, G., Hedrich, T., Gotman, J., Lina, J.-M., Grova, C., et al. (2016). Detection and magnetic source imaging of fast oscillations (40–160 Hz) recorded with magnetoencephalography in focal epilepsy patients. *Brain Topogr.* 29, 218–231. doi: 10.1007/s10548-016-0471-9
- Wang, S., Hu, L., and Cao, L. (2017). "Perceiving the next choice with comprehensive transaction embeddings for online recommendation," in *ECML/PKDD* (Springer), 285–302.
- Wu, T., Ge, S., Zhang, R., Liu, H., Chen, Q., Zhao, R., et al. (2014). Neuromagnetic coherence of epileptic activity: an meg study. *Seizure* 23, 417–423. doi: 10.1016/j.seizure.2014.01.022
- Xiang, J., Liu, Y., Wang, Y., Kirtman, E. G., Kotecha, R., Chen, Y., et al. (2009). Frequency and spatial characteristics of high-frequency neuromagnetic signals in childhood epilepsy. *Epileptic Disord.* 11, 113–125. doi: 10.1684/epd.2009.0253
- Xiang, J., Wang, Y., Chen, Y., Liu, Y., Kotecha, R., Huo, X., et al. (2010). Noninvasive localization of epileptogenic zones with ictal high-frequency neuromagnetic signals: case report. *J. Neurosurg. Pediatr.* 5, 113–122. doi: 10.3171/2009.8.PEDS09345
- Yamakawa, T., Miyajima, M., Fujiwara, K., Kano, M., Suzuki, Y., Watanabe, Y., et al. (2020). Wearable epileptic seizure prediction system with machine-learning-based anomaly detection of heart rate variability. *Sensors* 20, 3987. doi: 10.3390/s20143987
- Zelmann, R., Mari, F., Jacobs, J., Zijlmans, M., Chander, R., and Gotman, J. (2010). "Automatic detector of high frequency oscillations for human recordings with macroelectrodes," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (Buenos Aires: IEEE), 2329–2333.
- Zelmann, R., Mari, F., Jacobs, J., Zijlmans, M., Dubeau, F., and Gotman, J. (2012). A comparison between detectors of high frequency oscillations. *Clin. Neurophysiol.* 123, 106–116. doi: 10.1016/j.clinph.2011.06.006
- Zhang, J., Richardson, J. D., and Dunkley, B. T. (2020). Classifying post-traumatic stress disorder using the magnetoencephalographic connectome and machine learning. *Sci. Rep.* 10:5937. doi: 10.1038/s41598-020-62713-5
- Zhang, Y., Ai, Q., Chen, X., and Croft, W. B. (2017). "Joint representation learning for top-n recommendation with heterogeneous information sources," in *CIKM '17: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (Singapore: ACM), 1449–1458.
- Zhao, X., Niu, Z., and Chen, W. (2013). Interest before liking: two-step recommendation approaches. *Knowledge Based Syst.* 48, 46–56. doi: 10.1016/j.knosys.2013.04.009
- Zhao, X., Niu, Z., Chen, W., Shi, C., Niu, K., and Liu, D. (2014). A hybrid approach of topic model and matrix factorization based on two-step recommendation framework. *J. Intell. Inf. Syst.* 44, 335–353. doi: 10.1007/s10844-014-0334-3