# SAGA: SPARSE ADVERSARIAL ATTACK ON EEG-BASED BRAIN COMPUTER INTERFACE

*Boyuan Feng, Yuke Wang, Yufei Ding*

University of California, Santa Barbara

## ABSTRACT

With the recent advancement of the Brain-Computer Interface (BCI), Electroencephalogram (EEG) analytics gain a lot of research attention from various domains. Understanding the vulnerabilities of EEG analytics is important for safely applying this emerging technology in our daily life. Recent studies show that EEG analytics are vulnerable to adversarial attacks when adding small perturbations on the EEG data. However, fewer research efforts have been devoted to the robustness of EEG analytics under sparse perturbations that attack only small portions of the data. In this paper, we conduct the first in-depth study on the robustness of EEG analytics under sparse perturbations and propose the first Sparse Adversarial eeG Attack, SAGA, to identify weakness of EEG analytics. Specifically, by viewing EEG data as time series collected from several channels, we design an adaptive mask to uniformly represent diverse sparsity in adversarial attacks. We further introduce a PGD-based iterative solver to automatically select the time steps and channels under the given sparsity constraints and effectively identify the adversarial examples on EEG data. Extensive experiments show that SAGA can effectively generate sparse perturbations and introduces a $77.02\%$ accuracy drop on average by only perturbing $5\%$ channels and time steps.

## 1. INTRODUCTION

With the increasing popularity of Brain-Computer Interface (BCI), numerous research and industry efforts have been devoted to analyze the data and facilitate better designs. One of the most important designs is Electroencephalography (EEG) based BCI due to its convenience, which collects the brain electrical activity by attaching metal electrodes to the scalp. In particular, given $C$ electrodes and a time period $T$, it collects $C$-channel values at each time step and generate a $C \times T$ matrix as the EEG data. Recent studies [1–4] show that, when analyzing the EEG data with deep neural networks (DNN), we can usually achieve high accuracy across various BCI tasks, such as emotion recognition [5], intention recognition [6, 7], and epileptic seizure detection [8, 9].

Such wide applications of EEG analytics motivate the investigation of their robustness and reliability. One initial study [10] shows that adversarial attacks from the computer vision domain, such as FGSM [11] and iFGSM [12], can also dramatically change the outputs of EEG models by introducing perturbations on the EEG data. This approach assumes a strong attack capability that perturbs all time steps and all channels. However, it ignores the intrinsic channel interaction in EEG data that each channel represents a brain area and multiple brain areas interact to perform a specific task. It also ignores the importance of exploiting temporal propagation in EEG data, since EEG data from each channel across time steps can be viewed as time series.

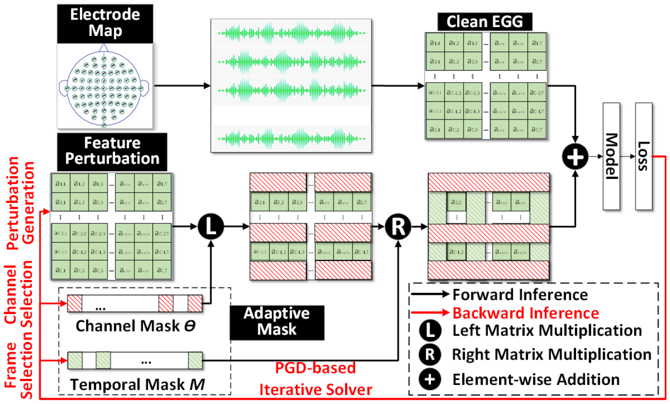
Electrode Map
Clean EGG
$a_{1,1}$ $a_{1,2}$ $a_{1,3}$ ... $a_{1,T}$
$a_{2,1}$ $a_{2,2}$ $a_{2,3}$ ... $a_{2,T}$
$a_{C-2,1}$ $a_{C-2,2}$ $a_{C-2,3}$ ... $a_{C-2,T}$
$a_{C-1,1}$ $a_{C-1,2}$ $a_{C-1,3}$ ... $a_{C-1,T}$
$a_{C,1}$ $a_{C,2}$ $a_{C,3}$ ... $a_{C,T}$
Feature Perturbation
Perturbation Generation
Channel Selection
Frame Selection
Channel Mask Θ
Temporal Mask M
Adaptive Mask
PGD-based Iterative Solver
L
R
+
Model
Loss
Forward Inference
Backward Inference
L Left Matrix Multiplication
R Right Matrix Multiplication
+ Element-wise Addition

**Fig. 1**: Overview of SAGA

In this paper, we focus on the sparse adversarial attack on EEG analytics with a weak assumption on attack capability that only sparsely attacks a small portion of channels and time steps. In practice, such sparse perturbations could come from the failure of a few channels or electrodes (*e.g.*, 3 out of 64 electrodes) at a small portion of time steps. However, it is a non-trivial task because of several challenges. First, we need a uniform optimization framework to formulate the attacking capability under various sparsity constraints. While existing works [13] usually rely on regularization techniques to implicitly introduce sparsity, this framework should be able to explicitly apply sparsity constraints on both the time- and channel-dimensions. Second, a sophisticated method is required to effectively solve the optimization problem under given sparsity constraints. Existing works with stochastic gradient descent (SGD) cannot solve the problem under constraints and usually require ad-hoc modifications [14, 15] on gradients to satisfy the constraints.

To tackle these challenges, we propose SAGA to enable the sparse adversarial attack on EEG analytics, as illustrated in Figure 1. First, we propose an *adaptive mask* on the time dimension and the channel dimension, which uniformly encodes diverse sparsity constraints. Second, we propose a projection gradient descent (PGD) based iterative solver to effectively generate adversarial examples while satisfying sparsity constraints with theoretical support. Extensive experiments on mainstream EEG datasets and models show that SAGA can effectively degrade accuracy by 77.02% under 5% sparsity constraints.

## 2. BACKGROUND AND RELATED WORK

**EEG-based Brain-Computer Interface.** EEG plays an important role in many Brain-Computer Interface (BCI) applications, such as emotion recognition [5] and epileptic seizure detection [8, 9]. EEG data is collected across a time period by a set of electrodes attached to the scalp. At each time step $t$ of a time period, $C$ electrodes collect $C$ scalar values to describe the human brain state

$$X_{1,t}, X_{2,t}, ..., X_{C,t}, \quad t \in \{1, 2, ..., T\} \tag{1}$$

During this time period, the person under test is asked to behave in a specific way (*e.g.*, moving hand or foot), which serves as the ground truth label $Y$.

In EEG analytics, we usually treat the collected EEG data $X \in \mathcal{R}^{C\times T}$ as features and build models to predict the behavior $Y$. Existing works on EEG analytics usually rely on manual feature extraction [16–18], leading to significant manual efforts. Recent studies [1–4] introduce deep learning into EEG analytics to automatically learn the features, showing significant improvement in accuracy. In this paper, we will provide a thorough study of the robustness of the deep-learning-based EEG analytics.

**Adversarial Attack.** Adversarial attack changes the deep learning prediction by adding small feature perturbations to the input data. It widely exists in various deep learning domains [11, 12, 19] (*e.g.*, computer vision and audio recognition). One popular attacking method on computer vision models is the fast gradient sign method (FGSM) [11] that utilizes gradients to generate feature perturbations. Another method, iFGSM [12], strengthens the adversarial attack by iteratively applying the FGSM. One recent study [10] has shown that these two methods can also attack deep learning models for EEG analytics. However, it assumes attacking all channels and all time steps simultaneously and cannot work effectively under sparsity constraints that attacking only a small portion of channels and time steps. In this paper, we propose SAGA, the first sparse adversarial attack on EEG analytics.

## 3. PROPOSED METHOD

We first define notations in this paper. We consider an input EEG data $X \in \mathcal{R}^{C\times T}$, where $C$ is the number of channels and $T$ is the length of the time period. A single ground truth label $Y$ exists during this time period, indicating a specific user behavior (e.g., moving hand or foot). Following the common practice in EEG analytics, there is a deep learning model $f_W(\cdot)$ that makes a prediction with the EEG data $X$ and the model parameter $W$. Here, $W$ is learned by minimizing the cross-entropy loss $\ell(X, Y)$.

### 3.1. Adaptive Mask

We propose an adaptive mask to uniformly formulate the EEG attack under diverse sparsity constraints. On the channel dimension, we propose the channel mask $\theta \in [0, 1]^C$ to encode two attack properties. First, a large magnitude of the value $\theta_c$ for channel $c$ indicates that channel $c$ plays an important role in attacking EEG analytics. Second, a zero value $\theta_c$ for a channel $c$ indicates that this channel will not involve in the adversarial attack under the sparsity constraints. Formally, given an un-constrained feature perturbation $X_\delta \in \mathcal{R}^{C\times T}$, we represent the constrained perturbation $\tilde{X} \in \mathcal{R}^{C\times T}$ under channel mask $\theta$ with left matrix multiplication

$$\tilde{X} = diag(\theta) \cdot X_\delta \tag{2}$$

Here, the left multiplication provides a row-wise selection in the channels according to the positions that $\theta_c > 0$. When generating adversarial perturbations, we apply a cardinality constraint $card(\theta) < \epsilon_\theta$ on the channel mask to explicitly encourage channel selection. Here, $card(\cdot)$ is the cardinality function that counts the number of non-zero values.

On the temporal dimension, we propose the temporal mask $M \in [0, 1]^T$ to select the time steps for the attack. Here, we also apply a cardinality constraint $card(M) < \epsilon_M$ on the temporal mask such that only a small portion of time steps are attacked. Different from the channel mask, we adopt the right matrix multiplication to provide a column-wise selection in the time steps according to the positions that $M_t > 0$

$$\tilde{X} = X_\delta \cdot diag(M) \tag{3}$$

In SAGA, we uniformly encode both sparsity constraints in the channel dimension and the temporal dimension as

$$\tilde{X} = diag(\theta) \cdot X_\delta \cdot diag(M) \tag{4}$$

Here, various sparsity constraints on the time steps and channels can be uniformly encoded by independently applying sparsity constraints on $\theta$ and $M$. From an element-wise perspective, we have

$$\tilde{X}_{ij} = \theta_i M_j X_{\delta,ij} \tag{5}$$

where a perturbation $X_{\delta,ij}$ at channel $i$ and time step $j$ is allowed only if both $\theta_i$ and $M_j$ are non-zero values.

With the adaptive mask, we formulate the sparse attack on EEG data as a non-convex optimization problem under constraints. Formally, during attack, we aim to solve the following optimization problem

$$\begin{aligned} \min_{\theta, M, X_\delta} \quad & -\ell(X + diag(\theta) \cdot X_\delta \cdot diag(M), Y) \\ s.t. \quad & card(\theta) \leq \epsilon_\theta, \;\; card(M) \leq \epsilon_M \end{aligned} \tag{6}$$

Here, $\theta \in \mathcal{R}^{C\times T}$ and $M \in \mathcal{R}^{C\times T}$ selects channels and time steps, respectively. $\epsilon_\theta$ and $\epsilon_M$ are user-specified attacking capability on the channel dimension and the temporal dimension. Our attack method generates sparse perturbations to degrade the model prediction accuracy while satisfying the given constraints of $\epsilon_\theta$ and $\epsilon_M$.

### 3.2. PGD-based Iterative Solver

While the adaptive mask effectively encodes the sparsity in channels and time steps, it is still challenging to solve the optimization problem 6. While existing approaches usually attack with stochastic gradient descent (SGD), it cannot be easily adapted to solve optimization problems under constraints without ad-hoc modifications on the gradients. In addition, the cardinality constraints in Equation 6 is non-differentiable, which makes it even harder to adopt SGD.

To this end, we propose a PGD-based iterative solver to effectively solve Equation 6, as summarized in Algorithm 1. It shows merits in automatically selecting the channel mask and the temporal mask under sparsity constraints while effectively generating adversarial examples under the selected masks. We first reformulate the equation 6 as

$$\begin{aligned} \min_{\theta, M, X_\delta} \quad & -\ell(X + diag(\theta) \cdot X_\delta \cdot diag(M), Y) \\ & + h_1(\theta) + h_2(M) \end{aligned} \tag{7}$$

where $h_1(\theta)$ and $h_2(M)$ are two indicator functions. In particular, $h_1(\theta) = 0$ if $card(\theta) \leq \epsilon_\theta$; $= \infty$, otherwise. Similarly, we have $h_2(M) = 0$ if $card(M) \leq \epsilon_M$; $= \infty$, otherwise. Then, we iteratively attack one of the variables while keeping other variables fixed.

At each iteration $k$, we first attack the channel mask $\theta$ while fixing the temporal mask $M$ and the unconstrained feature perturbation $X_\delta$

$$\begin{aligned} \theta^{(k+1)} &= \Pi_{S_1}[\theta^{(k)} + \eta_k g_\theta^{(k)}] \\ g_\theta^{(k)} &= \frac{\partial}{\partial \theta} \ell(Y, X + diag(\theta) X_\delta^{(k)} diag(M^{(k)})) \end{aligned} \tag{8}$$

Here, $S_1 = \{\theta \in [0,1]^C \mid card(\theta) \leq \epsilon_\theta\}$ and $\Pi_{S_1}(\cdot)$ is a projection operation. Generally, the projection operation is hard to compute and may involve an additional iterative procedure that is time-consuming. However, $S_1$ shows geometric properties that restrict the solution to grid points, which can be exploited to solve analytically. To this end, we introduce the cardinality projection [20] to efficiently compute the channel mask $\theta$. In particular, $\Pi_{S_1}$ ranks elements in $\theta^{(k)} + \eta_k g_\theta^{(k)}$ by its magnitude and keeps only the top-$\epsilon_\theta$ elements. Similarly, we can select the temporal mask with

$$\begin{aligned} M^{(k+1)} &= \Pi_{S_2}[M^{(k)} + \eta_k g_M^{(k)}] \\ g_M^{(k)} &= \frac{\partial}{\partial M} \ell(Y, X + diag(\theta^{(k+1)}) X_\delta^{(k)} diag(M)) \end{aligned} \tag{9}$$

Here, $S_2 = \{M \in [0,1]^T \mid card(M) \leq \epsilon_M\}$ and $\Pi_{S_2}(\cdot)$ is the cardinality projection operation. Once we have selected the channel mask $\theta^{(k+1)}$ and the temporal mask $M^{(k+1)}$, we update the unconstrained feature perturbation as

$$\begin{aligned} X_\delta^{(k+1)} &= X_\delta^{(k)} + \eta_k g_X^{(k)} \\ g_X^{(k)} &= \frac{\partial}{\partial X_\delta} \ell(Y, X + diag(\theta^{(k+1)}) X_\delta diag(M^{(k+1)})) \end{aligned} \tag{10}$$

**Algorithm 1:** Iterative Attack to solve Problem 6.

```
Input: Given X, fixed weight $W$, learning rate $\eta_k$, and
    iteration number $K$
Randomly initialize $\theta^{(1)}$, $M^{(1)}$, and $X_\delta^{(1)}$.
for $k = 1, 2, ..., K$ do
    Channel Selection on $\theta$:
    $\theta^{(k+1)} = \Pi_{S_1}[\theta^{(k)} + \eta_k g_\theta^{(k)}]$ with Eq 8.
    Frame Selection on $M$:
    $M^{(k+1)} = \Pi_{S_2}[M^{(k)} + \eta_k g_M^{(k)}]$ with Eq 9.
    Perturbation Generation on $X_\delta$:
    $X_\delta^{(k+1)} = X_\delta^{(k)} + \eta_k g_X^{(k)}$ with Eq 10.
end
Return channel mask $\theta^{(K)}$, frame mask $M^{(K)}$, and feature
    perturbation $X_\delta^{(K)}$.
```

## 4. EVALUATION

In this section, we evaluate SAGA on three EEG datasets and compare SAGA with two attack algorithms to show its effectiveness.

**Dataset.** We evaluate on three datasets from the EEG repository MNE [21] to cover the vast majority of EEG analytics. In particular, we utilize SPM dataset [22] on face perception, MI dataset [23, 24] on motor imagery, and ERP dataset [21] on visual-audio stimulus.

**Models.** We focus on two popular deep learning models from the open-source arl-eeg repository [25] – EEGNet [1] and DeepConv [2]. These two models have been widely utilized in analyzing visual stimulus, movement-related cortical potential, and sensorimotor rhythm. We use the default parameter setting in arl-eeg repository.

**Baselines and Metrics.** To evaluate the effectiveness of SAGA, we compare it with two state-of-the-art attack methods on EEG analytics [10] – FGSM [11] and iFGSM [12]. To enable these attack methods to generate adversarial examples under sparsity constraints, we uniformly sample the channels and time steps that allow the attack. By contrast, SAGA automatically selects channels and time steps to attack under the same sparsity constraints.

During the evaluation, we report the *accuracy* to show the effectiveness of adversarial attacks. To measure the perceptibility, we report the $L_2$ *norm* as a global measurement and the *distortion* as a local measurement. Formally, given the clean EEG data $X \in \mathcal{R}^{C \times T}$ and a sparse perturbation $\Delta \in \mathcal{R}^{C \times T}$, we have $norm = ||\Delta||_2$ and $Distortion = max_{c,t}|\Delta_{c,t}|/max_{c,t}|X_{c,t}|$. In addition, we report the *time sparsity* on the time dimension, measured by the number of

| Dataset | Method | EEGNet |  | DeepConv |  |
|---|---|---|---|---|---|
|  |  | Acc. (%) | Norm | Acc. (%) | Norm |
| SPM | Clean | 94.12 | - | 85.29 | - |
|  | FGSM | 70.59 | 112.9 | 61.76 | 112.9 |
|  | iFGSM | 70.59 | 112.8 | 58.82 | 112.7 |
|  | **SAGA** | 0 | 112.1 | 0 | 111.2 |
| MI | Clean | 99.4 | - | 66.67 | - |
|  | FGSM | 88.89 | 164.3 | 66.67 | 164.3 |
|  | iFGSM | 88.89 | 161.3 | 55.56 | 133.3 |
|  | **SAGA** | 11.11 | 155.2 | 11.11 | 132.3 |
| ERP | Clean | 91.67 | - | 72.22 | - |
|  | FGSM | 54.17 | 1.94 | 63.89 | 1.94 |
|  | iFGSM | 54.17 | 1.89 | 63.89 | 1.89 |
|  | **SAGA** | 16.67 | 1.86 | 8.33 | 1.84 |

**Table 1**: Overall performance under 5% sparsity constraint.

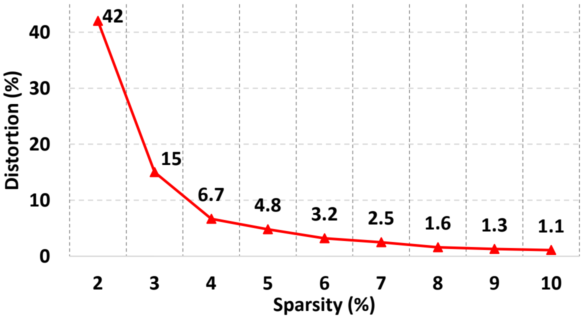

**Fig. 2**: Minimum distortion under diverse sparsity to fool EEGNet on SPM with 0% accuracy.

attacked time steps over the total number of time steps. Similarly, we have the *channel sparsity* on the channel dimension. Due to the page limit, we will focus on the same level of time sparsity and channel sparsity.

### 4.1. Overall Performance

Table 1 shows the overall performance comparison between SAGA and existing adversarial attacks under 5% sparsity constraint on both the time and channel dimensions. For a fair comparison, we use the same distortion for all methods. We observe that SAGA introduces 77.02% accuracy drop on average. Comparing with baseline methods, SAGA outperforms FGSM and iFGSM by 59.79% and 57.45%, respectively. The main benefit comes from the automatic selection of channels and time steps, indicating that different channels and time steps play different roles in EEG analytics. In addition, we observe that SAGA can usually achieve lower accuracy with a lower norm. This result shows that the adversarial examples from SAGA is less noticeable.

### 4.2. Ablation Study

In ablation study, we focus on EEGNet and SPM dataset due to the similar performance of SAGA across models and datasets.

**Minimum Distortion under Diverse Sparsity.** Figure 2 shows, under each sparsity, the minimum distortion required to fool EEGNet on SPM with 0% accuracy. Higher distortion usually leads to better attack effect, but also makes the adversarial example more perceptible. Overall, we observe that smaller distortions are required to fool the EEGNet on SPM as the sparsity increases. The reason is that it is easier to attack the model when perturbing more channels and time steps. We also observe that, for sparsity larger than 5%, a small distortion of less than 5% relative to the clean EEG data is sufficient to attack the model. This result shows that, even in generating sparse perturbations, we do not need to add large distortions on individual channels and time steps.

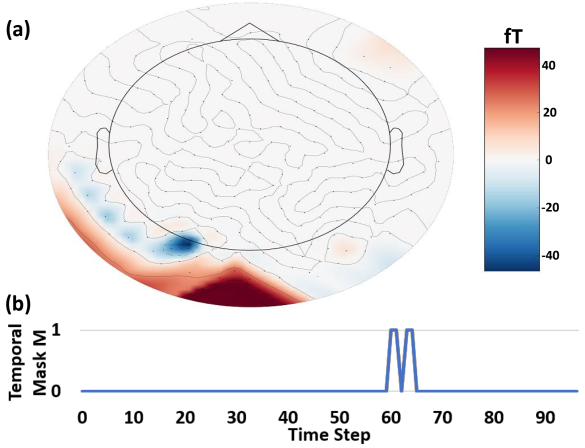

**Fig. 3**: Visualization of selected masks. (a) Channel mask. Blue and red area indicates the selected channels. (b) Temporal mask. One indicates selected time steps.

**Visualization of Temporal Mask and Channel Mask.** Figure 3 visualizes the selected time steps and channels under 5% sparsity. On the channel dimension, we observe that the selected channels clusters in a small brain area. This result shows that nearby brain areas usually collaborate to perform certain behaviors and sparse perturbations on a small area are sufficient for attacking EEG analytics. On the time dimension, we observe that SAGA attacks only on small contiguous time periods while not attacking on other time periods. The main reason is that EEG data can be essentially viewed as time series data and attacking at one time step can effectively propagate to other time steps.

## 5. CONCLUSION

This work focuses on the adversarial attack for EEG analytics under sparsity constraints. We propose SAGA, the first sparse attack to identify the weakness of the EEG analytics. In particular, SAGA utilizes an adaptive mask to uniformly formulate diverse sparsity constraints and a PGD-based iterative solver to automatically select the time and channels to attack. Extensive evaluations further highlight SAGA's advantage against state-of-the-art attacks on a large set of models and datasets.

## 6. ACKNOWLEDGEMENTS

We thank all anonymous reviewers for their valuable comments. This work was supported in part by NSF 1925717.

## 7. REFERENCES

[1] V. Lawhern, A. Solon, N. Waytowich, S. Gordon, C. Hung, and B. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of Neural Engineering*, 2018.

[2] S. Tibor, S. Tobias, F. Josef, G. Martin, E. Katharina, T. Michael, B. Wolfram, and B. Tonio, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human Brain Mapping*, vol. 38.

[3] M. Jalilpour-Monesi, B. Accou, J. Montoya-Martínez, T. Francart, and H. Hamme, "An LSTM based architecture to relate speech stimulus to eeg," in *ICASSP*, 2020.

[4] Y. Peng, Q. Li, W. Kong, J. Zhang, B. Lu, and A. Cichocki, "Joint semi-supervised feature auto-weighting and classification model for eeg-based cross-subject sleep quality evaluation," in *ICASSP*, 2020.

[5] T. Zhang, Z. Cui, C. Xu, W. Zheng, and J. Yang, "Variational pathway reasoning for EEG emotion recognition," in *AAAI*, 2020.

[6] D. Zhang, L. Yao, X. Zhang, S. Wang, W. Chen, and R. Boots, "Cascade and parallel convolutional recurrent neural networks on eeg-based intention recognition for brain computer interface.," in *AAAI*, 2018.

[7] Z. Fang, W. Wang, S. Ren, J. Wang, W. Shi, X. Liang, C. Fan, and Z. Hou, "Learning regional attention convolutional neural network for motion intention recognition based on eeg data," in *IJCAI*, 2020.

[8] X. Zhao, J. Solé-Casals, B. Li, Z. Huang, A. Wang, J. Cao, T. Tanaka, and Q. Zhao, "Classification of epileptic ieeg signals by cnn and data augmentation," in *ICASSP*, 2020.

[9] P. Boonyakitanont, A. Lek-uthai, and J. Songsiri, "Automatic epileptic seizure onset-offset detection based on cnn in scalp eeg," in *ICASSP*, 2020.

[10] L. Meng, C. Lin, T. Jung, and D. Wu, "White-box target attack for eeg-based bci regression problems," in *ICONIP*, 2019.

[11] I. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in *ICLR*, 2015.

[12] A. Kurakin, I. Goodfellow, and S. Bengio, "Adversarial machine learning at scale," *CoRR*, 2016.

[13] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistic Society, Series B*, vol. 58, pp. 267–288, 1994.

[14] C. Matthieu, B. Yoshua, and D. Jean-Pierre, "Binaryconnect: Training deep neural networks with binary weights during propagations," in *NeurIPS*. 2015.

[15] M. Rastegari, V. Ordonez, J. Redmon, and A. Farhadi, "Xnor-net: Imagenet classification using binary convolutional neural networks," in *ECCV*, 2016.

[16] A. Shoeb, A. Kharbouch, J. Soegaard, S. Schachter, and J. Guttag, "An algorithm for detecting seizure termination in scalp eeg," in *Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011.

[17] L. Orosco, A.G. Correa, P. Diez, and E. Laciar, "Patient non-specific algorithm for seizures detection in scalp eeg," in *Computers in Biology and Medicine*, 2016.

[18] G. Chandel, P. Upadhyaya, O. Farooq, and Y.U. Khan, "Detection of seizure event and its onset/offset using orthonormal triadic wavelet based features," in *IRBM*, 2019.

[19] X. Liu, K. Wan, Y. Ding, X. Zhang, and Q. Zhu, "Weighted-sampling audio adversarial example attack," in *AAAI*, 2020.

[20] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, 2011.

[21] A. Gramfort, M. Luessi, E. Larson, D. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. Hämäläinen, "Meg and eeg data analysis with mne-python," in *Frontiers in Neuroscience*, 2013.

[22] Rik Henson, "SPM Datasets: Multi-modal face dataset," `https://www.fil.ion.ucl.ac.uk/spm/data/mmfaces/`.

[23] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw., "Bci2000: a general-purpose brain-computer interface (bci) system.," in *IEEE Transactions on Biomedical Engineering*, 2004.

[24] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals.," in *Circulation*, 2000.

[25] Army Research Laboratory, "Arl eeg model project," `https://github.com/vlawhern/arl-eegmodels`.