IJCNN 2019. International Joint Conference on Neural Networks. Budapest, Hungary. 14-19 July 2019

Improved Techniques for Building EEG Feature Filters

Yue Yao
Australian National University
Canberra, Australia
yue.yao@anu.edu.au

Josephine Plested
Australian National University
Canberra, Australia
jo.plested@anu.edu.au

Tom Gedeon
Australian National University
Canberra, Australia
tom@cs.anu.edu.au

Yuchi Liu
Australian National University
Canberra, Australia
yuchi.liu@anu.edu.au

Zhengjie Wang
Australian National University
Canberra, Australia
u6259550@anu.edu.au

***Abstract*—Recent advances in the generative adversarial network (GAN) based image translation have shown its potential of being an image style transformer. Similarly, defined as a style transformer for physiological signals, a feature filter is used to filter privacy-related features while still keeping useful features. However, existing feature filter techniques have three problems: (1) the privacy-related features cannot be filtered out to the extent we need through a simple Conv-Deconv generator structure, and (2) the generator cannot control the semantics (maintain desired features) of given physiological signals. To address these problems, we utilize deeper neural networks and adopt techniques from domain adaptation. This includes semantic loss and a GAN based model structure with two generators, two discriminators and a classifier to form a game of five. Our results on the UCI EEG dataset demonstrate that our model can simultaneously (1) achieve the state-of-the-art accuracy removal for the privacy-related feature, (2) reduce the desired feature removal accuracy drop, and (3) make the filtered signals can be interpreted or visually checked.**

***Index Terms*—Deep Learning, EEG, Convolution Neural Network, Image Translation, Generative Adversarial Nets**

I. INTRODUCTION

Every single moment, people communicate with the world. Such communication between the brain and the external environment needs to be done through the peripheral nerves and muscle channels. However, the availability of the brain-computer Interface (BCI) provides such a non-muscle controlled communication channel, enabling the human brain to interact with the environment directly, using brain activity as a control interface [34], [15]. In real practice, the control interface is generally chosen from electroencephalograph (EEG) signals [19].

As an essential part of Brain-Computer Interfaces (BCIs), the EEG, also known as brainwaves, has found a variety of exciting and useful applications for users and has become increasingly important. Gathered from the scalp, the EEG is a signal containing information about the electrical activity of the brain. Electrodes placed on the scalp are used to capture electrical information from the brain under the scalp, bone and other tissues. Since it is an overall measurement of human brain electrical activity, it may contain a wealth of information. This is the reason why EEG can be applied to diverse areas like disease identification [30], personal recognition [28], visual image generation using brainwaves [21], and brain typing [20].

However, from the viewpoint of data analysis, automatic EEG analysis is challenging due to the inherent feature of bio-signals. Apart from the well-known low signal-noise ratio [4], [32], [26], data format varies [24], there is limited training data [12] and large individual difference [14], [25], [33]. One source of ambiguity is the fused nature of the features, which is common for most bio-signal feature learning tasks. The fusion here means that any one experimental trial of signals contains both desired features and privacy-related features for given tasks. Also, due to the lack of macroscopic knowledge of the mechanism of EEG activity, this fused feature problem in EEG is more serious than many other physiological signals.

In real-world situations, for EEG based applications, customers not only require accuracy for the brain-computer interface but also need a competent level of privacy and information safety [29]. But unfortunately, EEG data contain a messy, vibrant symphony of personal information, including one's individuality, learning capacity and emotional information. That is, all brain activity related features that are recorded will be uploaded and can be used for legal or potentially illegal objectives. Current research has tried to specify several standards for operating with EEG data to protect users' privacy but that has not solved the problem fundamentally [13], [23], [2].

To address the above issue, Yao proposed a feature filter for short-term EEG signals [36]. The essence of privacy problems comes from that data containing multiple labels' information, with at least one type of label related to our task as well as irrelevant labels that will cause privacy problems. In this paper they are referred to as desired and privacy-related features respectively. Shown in Fig. 1, it is based on the hypothesis that the existence of a certain feature for EEG signals can be

978-1-7281-2009-6/$31.00 ©2019 IEEE

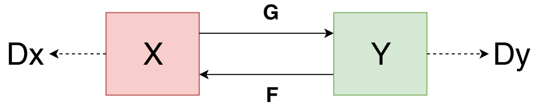

Fig. 1. Feature Filter Definition [36]. X is the original domain and Y is the target domain. The original domain in our definition refers to EEG singals with both desired features and privacy-related features. The target domain means EEG signals with desired features only.

defined precisely by the distribution X and Y. They described a feature filter as a task of domain translation. In Fig. 1, X means a domain with both desired features and privacy-related features. Y means a domain with desired features but with privacy-related features filtered out. So the feature filter is defined as a map from the X domain to the Y domain. To perform such a mapping function, they utilized a CycleGAN structure which could perform such domain transformation. But directly applying such a structure still faces two problems, (1) in testing there were still nearly 20% of EEG signals where the privacy-related feature was not filtered out, and (2) almost 10% of EEG signals lost their desired features when the feature filter was applied.

The first problem is that some EEG signals will still be recognized as having privacy-related features. We believe this is because the existing structure has not been able to capture the difference between the two distributions well enough. For the generator, the simplest solution is to make the model "deeper" to capture hierarchical information from one EEG image. So we applied the ResNet model from Johnson [11] with nine residual blocks. For the discriminator, we used the "patchGAN" classifier [37] which has a better fit with the cycle consistent loss. The combination of ResNet generator and patchGAN has shown success in many image translation tasks [16], [37], and our experiment further demonstrates that it also works as part of a feature filter.

The second problem comes from the important requirement about feature filtering in keeping the target feature. For the desired feature accuracy drop, we believe there does not exist enough regularization for the CycleGAN structure. That is because of their two-player formulation between generator and discriminator. Specifically, the training objective from the original cycleGAN does not involve any indicator for the magnitude of the undesirable reduction in the target feature. As a result, we take ideas from domain adaptation [9], to use one additional classifier to specify the feature we need to keep. The classifier will now work as a regularization term which helps the feature filter to maintain the information we want.

II. Related Work

The generative adversarial network (GAN) is a powerful framework that usually has two neural networks compete in a minimax game [7]. GAN variants have achieved success in the image generation area [22], [16], [37]. A GAN has two major networks, a generator, which learns the real data distribution and generates images, and a discriminator, which learns to identify the real image. Throughout this progress, both the generator and the discriminator strive to minimize their own costs and gradually approach their goal.

Image-to-image translation aims to learn the mapping between two different image distributions [10]. Systems for this task can capture the style difference between two image distributions and then translate a given image from source domain to target domain. Among existing models, The Cycle-Consistent Adversarial Networks(CycleGAN) [37] is a popular image-to-image translation network for unpaired images. It uses an autoencoder-like structure to overcome the challenge of pairing images.

Domain adaptation is a problem that given source data X, source label Z_S, and target data Y but without target labels, the objective is to learn a model that predicts the label for target data Y. Among existing domain adaptation models, the domain-adversarial neural network [1] (DANN) is the first that utilizes the adversarial mechanism for extracting domain invariant features. The subsequent adversarial discriminative domain adaptation (ADDA) [31] provides a simple but powerful framework for domain adaptation and many of the current methods can be seen as special cases of ADDA [9], [27], [17]. Among them, CYCADA [9] is the one that combine the ideas of CycleGAN and ADDA, forming the cornerstone for our current research on feature filters.

III. Methodology

A. UCI EEG Dataset

Using the same configuration as Yao's work [36], we use the multi-label UCI EEG dataset. It is an alcoholism dataset which contains 122 participants with 45 diagnosed as control and 77 as alcoholism, forming 3,819 trials of control EEG signals and 7,033 trials of alcoholism EEG signals. Each trial of EEG is also labeled with stimuli information. There are five types of stimuli which are images selected from the Snodgrass and Vanderwart picture set. That is we can both put one trail of EEG into a two-class alcoholism classification model as well as a five-class stimulus classification model. For the UCI dataset, each trial of EEG signal is in one-second length, sampled at 256 Hz using 64 electrodes cross the scalp. We split the source distribution within subjects, which is randomly split as 7:1:2 for training, validation and testing for each alcoholism subject. The target distribution is the whole data from control subjects.

For each trial of EEG signal, the EEG2Img technique [3] will be used to transfer wave like data to grid-like data. Given one trial of EEG signal, Fast Fourier Transform (FFT) is performed on the time series to transfer the time domain information to frequency domain information. Then with the help of the 3D electrode position, the frequency domain information on 3D can be plotted into 2D EEG images by polar projection. The previous work on EEG images has shown success on both short-term EEG and long-term EEG signals [35], [3], [18], demonstrating that it is an effective method for extracting features from EEG signals.

The motivation for this is very straight-forward. For the EEG2Img method, theoretically, we can adjust the size of the

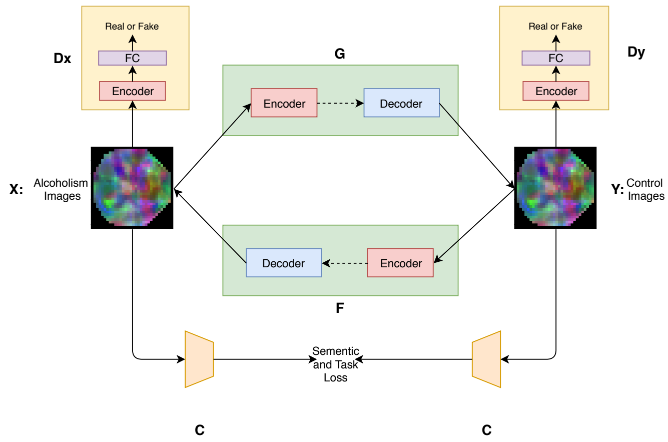

Fig. 2. Structure of Proposed Feature Filter

output EEG image as needed. So for one trial of EEG signal, we can directly transfer it to one EEG image with $32 \times 32 \times 3$ format which is a very typical format in the computer vision area and there are many mature and successful approaches and models for such formats. And since pictures are already generated by interpolation, $32 \times 32 \times 3$ format is enough to represent EEG frequency domain information. As a result, by utilizing this method, it is possible for us to adapt those computer vision models for EEG.

B. *Feature Filter for EEG*

For feature filter, we consider the problem of supervised domain transformation, where we are given source domain distribution X with both desired and privacy-related features, labels Z for desired features, target domain distribution Y with desired features only.

The objective of the feature filter is to directly learn a mapping from domain X to domain Y. So given an EEG image from X domain, the mapping representation in domain Y is our filter result. For this CycleGAN based Structure, the specific loss formulations are shown as follows.

Shown in Fig. 2, for UCI EEG dataset, the task of a feature filter map EEG images with the alcoholism condition to an EEG image with the control condition.

1) Loss Formulation: The objective of the feature filter is composed of three parts: adversarial loss, autoencoder loss and sentiment and classification loss. They can be expressed as:

A.Adversarial Loss:

The adversarial loss is the key part for the mapping from one distribution to another. For achieving this, the adversarial discriminator used to judge the following image is real or fake. For the loop $X \to G(X) \to F(G(X))$,

The ability to judge whether a image belongs to a certain distribution is given by the adversarial loss. For loop $X \to G(X) \to F(G(X))$, it is defined as:

$$L_{GAN}(G, D_Y, X, Y) = E_{x \sim pdata(x)}[log[(1 - D_Y(G(x))]] + E_{y \sim pdata(y)}[log D_Y(y)]$$

This is generally the standard format of GAN loss and used to make sure the generated samples are convincing. The adversarial loss for the loop $Y \to F(Y) \to G(F(Y))$ is in the similar format.

However, in real practice, the training of a GAN is quite unstable. Though the adversarial loss will force the generated image to look similar to real images, there is no guarantee for the direction of changes. To further make sure the feature filter meet our requirements, autoencoder loss and sentiment loss are introduced as regularization terms.

B.Autoencoder Loss:

The autoencoder loss is also named as reconstruction loss or cycle-consistency. It is basically an L1 loss which is used to keep $X \approx F(G(X))$, that is the generator will be forced to maintain features from the original image to have enough information to reconstruct the image during the backward loop. As a result, for loop $X \to G(X) \to F(G(X))$, it refers to:

$$L_{AL}(G, F) = E_{x \sim pdata(x)}[||F(G(x)) - x||_1]$$

Loop $Y \to F(Y) \to G(F(Y))$ has a similar autoencoder to make G(F(Y)) like Y.

C.Sentiment and Task Loss:

The sentiment and task loss originates from Hoffman's CYCADA model on domain adaptation [9]. Hoffman's solution is to train a cycleGAN model with sentiment and task loss to generate fake target data $fake_Y$ from source data X_S, thereby forming ($fake_Y$, Z_S) data label pairs.

Though the objective for domain adaptation is not related to our feature filter task, their proposed sentiment and task loss is useful for building a feature filter. In their proposed CYCADA model, the goal for using sentiment and task loss is to maintain labeled information when generating ($fake_{XT}$, Z_S) data label pairs. Such an idea satisfies the property that the desired features are maintained in our feature filter design.

The sentiment and task loss is given by an additional classifier C which gives labeled information. For the definition of task loss, it is basically a simple cross-entropy loss:

$$L_{task}(C, X, Z) = -E_{(x,z) \sim (X,Z)} \sum_{k=1}^{K} 1_{[k=z]} log(\sigma(C^{(k)}(x))),$$

where σ means the softmax function. In practice, the classifier will be trained on source domain X and desired label Z. As a result, loss $L_{task}(C, X, Z)$ will be used to show that the target feature label is retained.

So the classifier C works as a constraint by giving a semantically consistent loss. The semantic consistent loss will not take any explicit labeled information but focuses on the label consistency. That is the two generators will not change the labeled information when performing image translation. If we define $p(C, X) = argmax(C(X))$, the semantic consistency loss is as follows:

$$\begin{aligned} L_{sem}(G, F, X, Y, C) = L_{task}(C, F(Y), p(C, Y)) \\ + L_{task}(C, G(X), p(C, X)) \end{aligned}$$

As a conclusion, using the full loss functions mentioned above, we add those loss functions, and we have the final objective:

$$\begin{aligned} L_{total} &= L_{task}(C, X, Z) \\ &+ L_{GAN}(G, D_Y, X, Y) + L_{GAN}(G, D_X, Y, X) \\ &+ L_{AL}(G, F) + L_{AL}(F, G) \\ &+ L_{sem}(G, F, X, Y, C) \end{aligned}$$

2) Network Architecture: We first reconstruct Yao's work on feature filters to use a modified version of Image-wise Autoencoder as our generator (Shown in Table. 1.), and our discriminator is the combination of Image-wise Autoencoder and one fully connected layer works.

TABLE I
THE CONV-DECONV GENERATOR STRUCTURE [36]

Encoder	Decoder
Input $32 \times 32 \times 3$ Color Image	Input $128 \times 8 \times 8$ Matrix
4×4 conv, Leaky ReLU,	4×4 Deconv, Leaky ReLU,
4×4 conv, Leaky ReLU,	4×4 Deconv, Leaky ReLU,
3×3 conv, Leaky ReLU,	Tanh
3×3 conv, Leaky ReLU,	

For improving performance, we tried the ResNet-9 generator and patchGAN combination for training. The combination of ResNet generator and patchGAN achieves the best performance in many image translation applications [37]. Shown in Fig. 3, the residual-based generator is based on Johnson's ResNet model on super-resolution [11]. Similar to their work, our network is composed of one encoding block, nine residual blocks, and one decoding block. Each encoding or decoding block follows the two-stride convolution/deconvolution-InstanceNormReLU structure, and each residual block follows the convolution-InstanceNorm-ReLU-convolution-InstanceNorm residual connection structure. The advantage of using ResNet-9 is because it is capable of identifying the highways and produces straighter street blocks in the map, thereby making it easier for the generator to learn the identity function [8].

The patchGAN discriminator is derived from pix2pix [10], which is a paired image translation framework. The ordinary discriminator determines whether an image is real or fake from the entire image while the PatchGAN discriminator use local patches. For loop $X \rightarrow G(X) \rightarrow F(G(X))$, The discriminator D_y takes in two images, the real image Y and the generated image G(X), passes them through 5 downsampling convolutional-BatchNorm-LeakyReLU layers, and outputs a matrix for further classification. That is each element in the matrix corresponds to the classification of one patch. The advantage of using patchGAN is to avoid conflict with the autoencoder loss. Since we are using the final matrix to classify the image as real or fake, the patchGAN structure is used primarily to model high-frequency structure, whereas the autoencoder loss already provides low-frequency information [10].

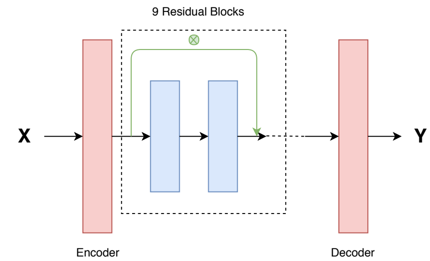

Fig. 3. Resnet-9 Generator Structure

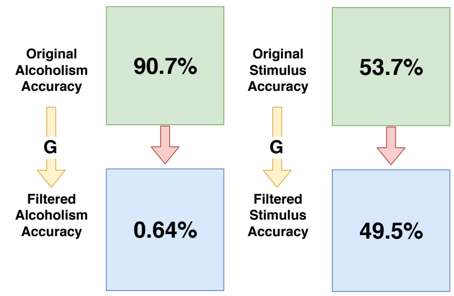

Fig. 4. Feature Filter Performance (alcoholism = 2 class, stimulus = 5 class)

C. Evaluation Method

We use the same configuration as Yao's work [36] to allow direct comparison. This consists of a pre-trained additional classifier to judge whether the features are maintained or not, and the additional classifiers for disease and stimulus are taken from the Image-wise autoencoder [35] and trained separately from the feature filter training. The additional classifiers are also trained in the within-subject setting in both alcoholism subjects and control subjects. The objective of the feature filter is to filter out privacy-related features while keeping desired features, as a result, it will be best if we can witness a significant alcoholism accuracy reduction with at most mild stimulus accuracy reduction.

IV. RESULTS AND DISCUSSION

Fig. 5. shows the visual example of the result of the feature filter. The left two columns map disease EEG images to control EEG images, the right two columns map control EEG images to the disease EEG images. From each direction, it can be seen that our feature filter has made a slight style transformation to images. However, those style changes are

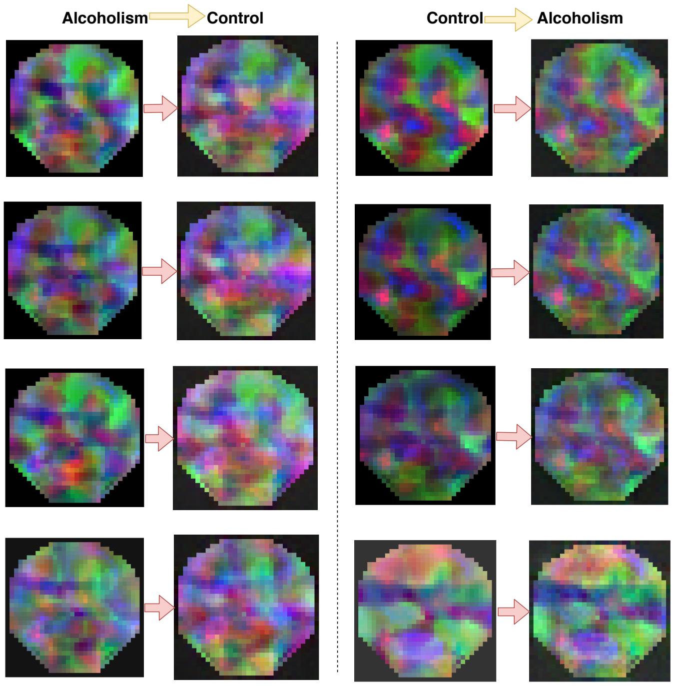

Fig. 5. Feature Filter Output

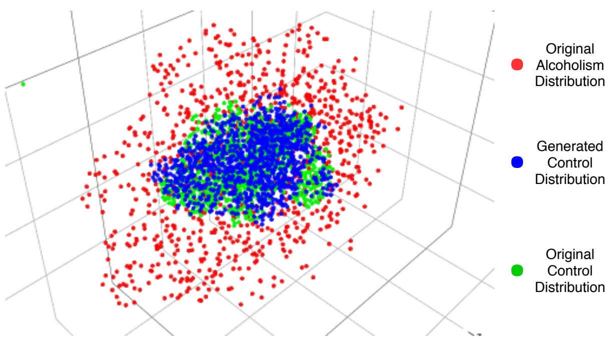

Fig. 6. Performance with t-SNE visualization

not interpretable since features from the original EEG images are not interpretable. But from t-SNE visualization in Fig. 6, we can see that the generated control image distribution is close to the original control distribution. Also, they have clear differences from the original alcoholism image distribution. Furthermore, from Fig. 4, initially, 90.7% of the original images are correctly classified as alcoholism. After our feature filter, only 0.6% of the images are classified as alcoholism. That is nearly all images have had their alcoholism information filtered out. At the same time, stimulus accuracy has only lost 4.2%, and the remaining accuracy is still well above chance since it is a 5-class classification problem for classifying stimuli.

Furthermore, one testing technique is to go through the feature filter multiple times. This idea is inspired by Ge's work for grammar error correction [5]: in their proposed work, they observed that some sentence with multiple grammatical

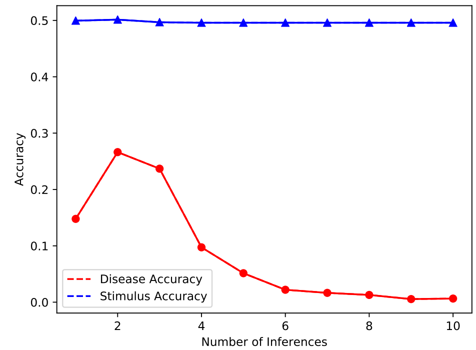

Fig. 7. Performance with the Number of Inferences

TABLE II
COMPARISON MODELS AND LOSS FUNCTIONS

Method	Alcoholism Acc %	Stimulus Acc %
G:Conv-Deconv D:Conv ($L_{GAN} + L_{AL}$) [36]	18.2	47.7
G:Resnet D:PatchGAN ($L_{GAN} + L_{AL}$)	0.643	48.9
G:Resnet D:PatchGAN (L_{total})	**0.642**	**49.5**

errors cannot be corrected by the Seq2Seq [6] inference using a single round of inference. So they involve multiple rounds of inference in both training and testing. In our work, we have not involved multiple inferences in training but merely used our trained feature filter to make multiple inferences on validation and test data. The result shown in Fig. 7 indicates that result is stable after six round of inference. The accuracy increases in the first 3 rounds, we think that is because our feature filter removes unstable factors rather than filtering out the privacy-related information in the first three rounds, but that need a further analysis to determine.

The performance difference between models and loss functions are shown in Table. 2. The result show that the best performance after multiple times of inference on the test set. We can see that Resnet and patchGAN contribute most to the performance boost. The sentiment loss and task loss contributes but does not achieve significant improvement on the drop in alcoholism accuracy. One hypothesis we have is that the stimulus classifier is currently far from a strong classifier. Our 53.7% is reasonable where chance is 20%, but cannot really be called a strong classifier. Thus, we think that could be one factor why adding sentiment and task loss has not achieved a larger improvement.

V. CONCLUSION

Building a feature filter will have a significant improvement on people's privacy protection. Previous work on feature filters still have problems in filtering out all privacy-related feature information and keep wanted feature loss in a reasonable range. This paper further improves the performance of the feature filter and nearly removed all privacy-related features by introducing deeper networks and semantic loss. The experiment results using accuracy drops show that our proposed feature filter can filter out a nearly all privacy-related features and still keep most of the desired features.

VI. Limitation and Future Work

The first limitation is that our method is based on EEG2Img and image translation techniques, which means that it is only suitable for short-term EEG signals. The design of a feature filter for long-term EEG signals remains to be solved. The second limitation is future work for the generator. The U-net structure is also applicable as a generator since it is also the current state of the art method for several image translation tasks. The third limitation is in our model we simply stack error functions but do not really optimize the training procedure. For further limiting the loss of wanted features, we can begin with the modification of training procedure for a GAN. Finally, we will include more experiments on different short-term EEG datasets in the future to test existing hypothesis on multi-inference technique and semantic loss.

References

[1] H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, and M. Marchand. Domain-adversarial neural networks. *arXiv preprint arXiv:1412.4446*, 2014.

[2] S. Al-Janabi, I. Al-Shourbaji, M. Shojafar, and S. Shamshirband. Survey of main challenges (security and privacy) in wireless body area networks for healthcare applications. *Egyptian Informatics Journal*, 18(2):113–122, 2017.

[3] P. Bashivan, I. Rish, M. Yeasin, and N. Codella. Learning representations from eeg with deep recurrent-convolutional neural networks. *arXiv preprint arXiv:1511.06448*, 2015.

[4] S. Gandhi, T. Oates, T. Mohsenin, and D. Hairston. Denoising time series data using asymmetric generative adversarial networks. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 285–296. Springer, 2018.

[5] T. Ge, F. Wei, and M. Zhou. Reaching human-level performance in automatic grammatical error correction: An empirical study. *arXiv preprint arXiv:1807.01270*, 2018.

[6] J. Gehring, M. Auli, D. Grangier, D. Yarats, and Y. N. Dauphin. Convolutional sequence to sequence learning. *arXiv preprint arXiv:1705.03122*, 2017.

[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[8] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[9] J. Hoffman, E. Tzeng, T. Park, J.-Y. Zhu, P. Isola, K. Saenko, A. A. Efros, and T. Darrell. Cycada: Cycle-consistent adversarial domain adaptation. *arXiv preprint arXiv:1711.03213*, 2017.

[10] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. *arXiv preprint*, 2017.

[11] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*, pages 694–711. Springer, 2016.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[13] J. Kulynych. Legal and ethical issues in neuroimaging research: human subjects protection, medical privacy, and the public communication of research results. *Brain and cognition*, 50(3):345–357, 2002.

[14] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance. Eegnet: A compact convolutional network for eeg-based brain-computer interfaces. *arXiv preprint arXiv:1611.08024*, 2016.

[15] M. A. Lebedev and M. A. Nicolelis. Brain–machine interfaces: past, present and future. *TRENDS in Neurosciences*, 29(9):536–546, 2006.

[16] M.-Y. Liu, T. Breuel, and J. Kautz. Unsupervised image-to-image translation networks. In *Advances in Neural Information Processing Systems*, pages 700–708, 2017.

[17] M.-Y. Liu and O. Tuzel. Coupled generative adversarial networks. In *Advances in neural information processing systems*, pages 469–477, 2016.

[18] Y. Liu, Y. Yao, W. Zhengjie, J. Plested, and T. Gedeon. Generalized alignment for multimodal physiological signal learning. In *Submited to 2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.

[19] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4(2):R1, 2007.

[20] U. Orhan, K. E. Hild, D. E. II, B. Roark, B. Oken, and M. Fried-Oken. Rsvp keyboard: An eeg based typing interface. In *Proceedings of the... IEEE International Conference on Acoustics, Speech, and Signal Processing/sponsored by the Institute of Electrical and Electronics Engineers Signal Processing Society. ICASSP*. NIH Public Access, 2012.

[21] S. Palazzo, C. Spampinato, I. Kavasidis, D. Giordano, and M. Shah. Generative adversarial networks conditioned by brain signals. *PDF available on ucf. edu*, 2017.

[22] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

[23] V. Robinson and E. B. Varghese. A novel approach for ensuring the privacy of eeg signals using application-specific feature extraction and aes algorithm. In *Inventive Computation Technologies (ICICT), International Conference on*, volume 2, pages 1–6. IEEE, 2016.

[24] A. Schlögl. An overview on data formats for biomedical signals. In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, pages 1557–1560. Springer, 2009.

[25] M. Schröder, T. N. Lal, T. Hinterberger, M. Bogdan, N. J. Hill, N. Birbaumer, W. Rosenstiel, and B. Schölkopf. Robust eeg channel selection across subjects for brain-computer interfaces. *EURASIP Journal on Applied Signal Processing*, 2005:3103–3112, 2005.

[26] S.-Y. Shao, K.-Q. Shen, C. J. Ong, E. P. Wilder-Smith, and X.-P. Li. Automatic eeg artifact removal: a weighted support vector machine approach with error correction. *IEEE Transactions on Biomedical Engineering*, 56(2):336–344, 2009.

[27] J. Shen, Y. Qu, W. Zhang, and Y. Yu. Adversarial representation learning for domain adaptation. *stat*, 1050:5, 2017.

[28] F. Su, L. Xia, A. Cai, Y. Wu, and J. Ma. Eeg-based personal identification: from proof-of-concept to a practical system. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 3728–3731. IEEE, 2010.

[29] K. Sundararajan. *Privacy and security issues in Brain Computer Interfaces*. PhD thesis, Auckland University of Technology, 2018.

[30] N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, and O. Kavehei. A generalised seizure prediction with convolutional neural networks for intracranial and scalp electroencephalogram data analysis. *arXiv preprint arXiv:1707.01976*, 2017.

[31] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7167–7176, 2017.

[32] J. A. Urigüen and B. Garcia-Zapirain. Eeg artifact removalstate-of-the-art and guidelines. *Journal of neural engineering*, 12(3):031001, 2015.

[33] Z. Wang, R. M. Hope, Z. Wang, Q. Ji, and W. D. Gray. Cross-subject workload classification with a hierarchical bayes model. *NeuroImage*, 59(1):64–69, 2012.

[34] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002.

[35] Y. Yao, J. Plested, and T. Gedeon. Deep feature learning and visualization for eeg recording using autoencoders. In *International Conference on Neural Information Processing (ICONIP) 2018*, page 13, 2018.

[36] Y. Yao, J. Plested, and T. Gedeon. A feature filter for eeg using gan-based autoencoder. In *International Conference on Neural Information Processing (ICONIP) 2018*, page 9, 2018.

[37] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv preprint*, 2017.