Image Reconstruction from Brain EEG Signals

KAIXUAN ZHANG, Northwestern University, USA

ACM Reference Format:
Kaixuan Zhang. 2023. Image Reconstruction from Brain EEG Signals. *ACM Trans. Graph.* 42, 4 (August 2023), 3 pages. https://doi.org/10.1145/3592127

1 INTRODUCTION

Significant progress has been achieved in the field of brain decoding, enabling the classification of image categories by analyzing neural activity. Subsequently, a growing body of research has focused on reconstructing the actual image from these neural signals. Typically, the reconstruction requires retinotopically organized neural data with high spatial resolution, such as fMRI signals [Shen et al. 2019; Stewart et al. 2014]. Although the effectiveness of decoding image from fMRI data has been proved, the practical applications of this approach are constrained by its high costs and potential invasiveness. In order to overcome these limitations, a number of studies have turned to electroencephalography (EEG) as an alternative, offering a convenient, cost-effective, and non-invasive method for gathering brain activity data.

In this report, we will summary a few papers regarding image decoding from brain EEG signals. The image decoding pipeline is given in Fig 1.

2 BACKGROUND

2.1 EEG Signals

Electroencephalography (EEG) signals serve as distinctive patterns of neural activity, representing the cumulative effect of active potentials originating from different regions of the brain at various time points. These signals exhibit varying latencies and involve populations of neurons active at each specific moment. An example of EEG signal is given Fig 1.

2.2 LSTM

Long Short-Term Memory (LSTM) [Hochreiter and Schmidhuber 1997] is a recurrent neural network that has been widely used in processing sequential data. Unlike feedforward neural networks, LSTM has feedback connections and the input and output of LSTM are both sequential data. LSTM has been applied to tasks such as speech recognition, machine translation, speech activity detection, robot control, video games, and healthcare.

Author's address: Kaixuan Zhang, kaixuanzhang2019@u.northwestern.edu, Northwestern University, Evanston, Illinois, USA, 60201.

0730-0301/2023/8-ART $15.00
https://doi.org/10.1145/3592127

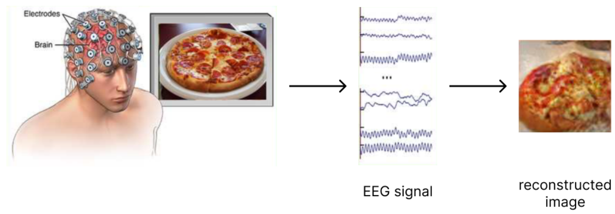

Fig. 1. Image reconstructino pipeline from brain EEG signals. Image is stolen from [Palazzo et al. 2017]

2.3 GAN and Conditional GAN

Generative Adversarial Networks (GAN) [Goodfellow et al. 2014] is a generative framework which includes two parts: a generative model G that captures the data distribution, and a discriminative model D that estimates the probability that a sample came from the training data rather than G.

A conditional GAN [Mirza and Osindero 2014] is a GAN whose generator G has two input: an image and also some extra information, such as a label for the image. The generated results of G can be conditioned on the extra information.

2.4 VAE

Variational AutoEncoder (VAE) is a also generative model. It has two parts: an encoder which maps the input into a latent distribution and a decoder which decode a vector in the latent distribution back to an image.

3 BRAIN IMAGE DECODING WITH MVAE

[Wakita et al. 2021] proposed an Multimodal Variational Auto Encoder (MVAE) based approach for reconstructing texture images from EEG signals. The database they adopted is [Orima and Motoyoshi 2021] which contains EEG signals for 166 natural texture images, with each signal measured for a period of 500 ms, 24 times, for each of 15 human observers.

In this study, the authors applied an extended method for the Multimodal Variational Auto Encoder (MVAE) [Wu and Goodman 2018], which allows inference of latent variables even under the partial observation of multimodal information aiming at reconstructing texture images only from EEG signals. The model MVAE includes a VAE for the texture image, and a VAE for the EGG signal. Both VAEs have the same latent space. During the training phase, the model can be trained with either one modality, or both. A figure of the pipeline is given in Fig 2

In order to evaluate the model, authors collected data through subjective test, which asks 6 observers to select the correct reconstructed texture among several other reconstructed textures, given the original texture. The reconstructed textures are generated by models trained with 10 fold cross validation. Experiment results

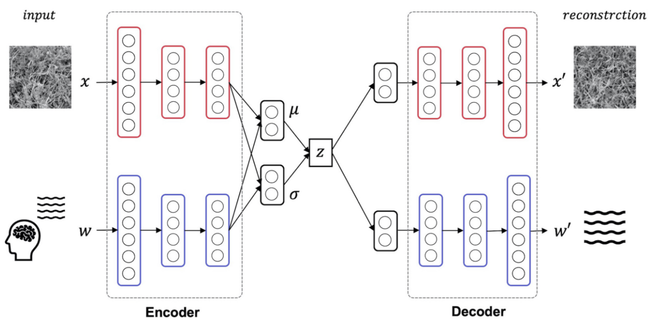

Fig. 2. Pipeline for [Wakita et al. 2021]. Image is stolen from their paper.

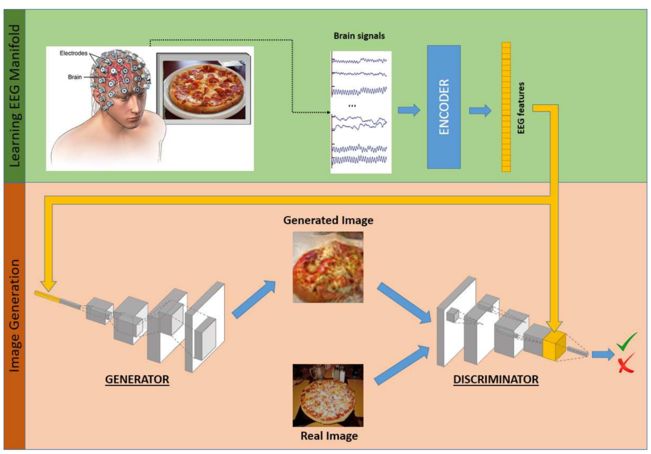

Fig. 3. Pipeline for [Palazzo et al. 2017]. Image is stolen from their paper.

show that the correct identification rate is about 70%, suggesting the reconstruction is successful. However, the authors also pointed out that the latent space is not rich enough, and more data would be helpful.

4 BRAIN IMAGE DECODING WITH CONDITIONAL GAN

[Palazzo et al. 2017] also proposed a framework, as shown in Fig 3, to decode image from EEG signal. The EEG signal is first transform to low dimensional vector by a LSTM, and the conditional GAN is trained with the ground truth image as well as the output from LSTM.

The authors collected a new dataset, with 50 images from 40 different object classes. Sixs participants are shown with all the 2,000 images. Each image is presented for 25 seconds followed with a 10 seconds pause.

The generated results are mainly evaluated by human. In [Palazzo et al. 2017], the authors show a few good results (Fig 4) and bad results (Fig 5), indicating that the proposed framework can successfully decode a few images from brain EEG signals but the model performance still have room to improve.

5 ETHICAL QUESTIONS

A common ethical question in AI research is how representative is the participants. Collecting EEG data requires a lab environments and usually the number of participants is not large, therefore guarantee the participants are representative is an important questions for researches that use EEG signals. A similar question is how representative is the data.

Another potential ethical question is privacy. Assuming the technology will be improved significantly, and in a few years we can easily decode images from brain EEG signals. Then what if we unintentionally decode information that a person do not want to share? Such as a painful childhood experience.

REFERENCES

Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Networks. arXiv:1406.2661 [stat.ML]

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-term Memory. *Neural computation* 9 (12 1997), 1735–80. https://doi.org/10.1162/neco.1997.9.8.1735

Mehdi Mirza and Simon Osindero. 2014. Conditional Generative Adversarial Nets. arXiv:1411.1784 [cs.LG]

Taiki Orima and Isamu Motoyoshi. 2021. Analysis and synthesis of natural texture perception from visual evoked potentials. *Frontiers in Neuroscience* (2021), 876.

S. Palazzo, C. Spampinato, I. Kavasidis, D. Giordano, and M. Shah. 2017. Generative Adversarial Networks Conditioned by Brain Signals. In *2017 IEEE International Conference on Computer Vision (ICCV)*. 3430–3438. https://doi.org/10.1109/ICCV.2017.369

Guohua Shen, Kshitij Dwivedi, Kei Majima, Tomoyasu Horikawa, and Yukiyasu Kamitani. 2019. End-to-end deep image reconstruction from human brain activity. *Frontiers in computational neuroscience* 13 (2019), 21.

Andrew X. Stewart, Antje Nuthmann, and Guido Sanguinetti. 2014. Single-trial classification of EEG in a visual object task using ICA and machine learning. *Journal of Neuroscience Methods* 228 (2014), 1–14. https://doi.org/10.1016/j.jneumeth.2014.02.014

Suguru Wakita, Taiki Orima, and Isamu Motoyoshi. 2021. Photorealistic reconstruction of visual texture from EEG signals. *Frontiers in Computational Neuroscience* 15 (2021), 754587.

Mike Wu and Noah Goodman. 2018. Multimodal generative models for scalable weakly-supervised learning. *Advances in neural information processing systems* 31 (2018).

(a) Airliner (b) Jack-o'-Lantern (c) Panda

Fig. 4. Good results from [Palazzo et al. 2017].

(a) Banana (b) Capuchin (c) Bolete

Fig. 5. Bad results from [Palazzo et al. 2017].