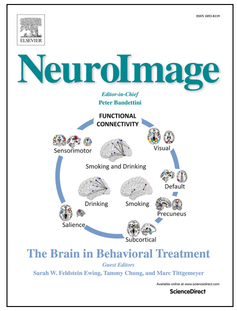
Accepted Manuscript



Generative adversarial networks for reconstructing natural images from brain activity

K. Seeliger, U. Güçlü, L. Ambrogioni, Y. Güçlütürk, M.A.J. van Gerven

PII: S1053-8119(18)30658-X

DOI: 10.1016/j.neuroimage.2018.07.043

Reference: YNIMG 15134

To appear in: *NeuroImage*

Received Date: 14 December 2017

Revised Date: 22 June 2018

Accepted Date: 16 July 2018

Please cite this article as: Seeliger, K., Güçlü, U., Ambrogioni, L., Güçlütürk, Y., van Gerven, M.A.J., Generative adversarial networks for reconstructing natural images from brain activity, *NeuroImage* (2018), doi: 10.1016/j.neuroimage.2018.07.043.

Generative adversarial networks for reconstructing natural images from brain activity

K. Seeliger[*], U. Güçlü, L. Ambrogioni, Y. Güçlütürk, M. A. J. van Gerven

Keywords
Vision, reconstruction, generative adversarial networks, fMRI

Abstract

We explore a method for reconstructing visual stimuli from brain activity. Using large databases of natural images we trained a deep convolutional generative adversarial network capable of generating gray scale photos, similar to stimuli presented during two functional magnetic resonance imaging experiments. Using a linear model we learned to predict the generative model's latent space from measured brain activity. The objective was to create an image similar to the presented stimulus image through the previously trained generator. Using this approach we were able to reconstruct structural and some semantic features of a proportion of the natural images sets. A behavioral test showed that subjects were capable of identifying a reconstruction of the original stimulus in 67.2% and 66.4% of the cases in a pairwise comparison for the two natural image datasets respectively. Our approach does not require end-to-end training of a large generative model on limited neuroimaging data. Rapid advances in generative modeling promise further improvements in reconstruction performance.

1 Introduction

Since the advent of functional magnetic resonance imaging (fMRI), numerous new research directions that leverage its exceptional spatial resolution, leading to classifiable brain activity patterns, have been explored (Haynes, 2015). New approaches to decoding specific brain states have demonstrated the benefits of pattern-based fMRI analysis. Pattern-based decoding from the visual system has shown that it is possible to decode edge orientation (Kamitani and Tong, 2005), perceived categories of both static and dynamic stimuli (Haxby, 2001; Huth et al., 2016), up to identifying a specific stimulus image (Kay et al., 2008) and generically identifying new categories from image descriptors predicted from brain activity (Horikawa and Kamitani, 2017).

Here we focus on an advanced problem in brain decoding, which is reconstructing a perceived (natural) visual stimulus. The reconstruction problem is demanding since the set of possible stimuli is effectively infinite. A functioning reconstruction system may however prove highly useful for neuroscience, for instance for studying synesthesia and optical illusions; or drive explorative insight into visual cortex activity when controlled experimental setups are difficult – such as during imagery or visual hallucinations. This problem has been explored at different spatial scales (e.g. invasively at the cellular level (Chang and Tsao, 2017)) and in different regions of the visual system (e.g. in the LGN (Stanley et al., 1999) and in the retina (Parthasarathy et al., 2017)). In this manuscript we discuss a new method for reconstruction from brain activity measured with fMRI. This approach was pioneered by Thirion et al. (2006), who reconstructed dot patterns with rotating Gabors from perception and imagery. Miyawaki et al. (2008) used binary 10×10 images as stimuli and demonstrated the possibility of decoding pixels independently from each other, reconstructing arbitrary new images with this basis set. Naselaris et al. (2009) introduced a combination of encoding brain activity with structural and semantic features, as well as a Bayesian framework to identify the most likely stimulus image from a very large image database given the

[*]Corresponding author: ✉ kseeliger@posteo.jp