Understanding illusory face perception in the human brain

Susan G Wardle (susan.wardle@nih.gov)
Laboratory of Brain & Cognition, NIMH/NIH
10 Center Dr, Bethesda, MD 20814 USA

Jessica Taubert (jessica.taubert@nih.gov)
Laboratory of Brain & Cognition, NIMH/NIH
10 Center Dr, Bethesda, MD 20814 USA

Lina Teichmann (lina.teichmann@mq.edu.au)
Department of Cognitive Science, Macquarie University
Sydney, NSW 2109 Australia

Chris I Baker (bakerchris@mail.nih.gov)
Laboratory of Brain & Cognition, NIMH/NIH
10 Center Dr, Bethesda, MD 20814 USA

Abstract:

Face pareidolia is the spontaneous misperception of illusory faces in inanimate objects. As a natural error of face detection, investigation of face pareidolia has potential to reveal new insight into the mechanisms underlying face and object recognition in the primate brain. To understand the temporal dynamics of illusory face perception, here we used magnetoencephalography (MEG) to measure the brain activation patterns of 22 human participants in response to 32 illusory faces in inanimate objects, 32 similar non-face objects, and 32 human faces. We compared the time-varying brain activation patterns to the representations of the stimuli obtained with common pre-trained deep neural networks and computational models of visual saliency. Whereas the saliency models showed an early and limited correlation with the MEG data peaking at ~100ms after stimulus onset, the deep neural networks showed a greater and sustained correlation across the time-course which generally increased with layer depth. However, both forms of computational model were outperformed by a simple model which grouped stimuli by category (faces versus objects). Together the results highlight that the dynamic response of the human brain to complex natural stimuli is only partially captured by existing computational models.

Keywords: face perception; object recognition; MEG; MVPA

Introduction

A common human experience is face pareidolia: the spontaneous misperception of illusory faces in inanimate objects such as fruit, trees, and the well-known face of the “Man in the Moon”. Pareidolia is a natural error of face detection and has recently been demonstrated in the rhesus macaque monkey (Taubert et al., 2017), suggesting that it is a fundamental aspect of the primate face detection system, rather than a uniquely human experience.

Face pareidolia presents an interesting example for the study of the neural mechanisms underlying face and object recognition in the human brain because (1) these stimuli have a simultaneous dual identity (object, face) and (2) face perception is divorced from the typical visual features that define a face. The presence of an illusory face in an object can be decoded from BOLD activation patterns measured with fMRI in both face and object-selective regions in the ventral stream (Wardle et al., 2017), however it is not yet understood how the temporal dynamics of illusory face perception compare to that for real faces.

Here we used MEG to characterize the temporal dynamics of spontaneous illusory face perception in the human brain. Further, we compared the brain activation patterns elicited by illusory faces with the output of computational models of visual saliency and deep neural networks (DNNs).

Methods

The visual stimuli were 96 photographs including 32 illusory faces in inanimate objects, 32 similar non-face objects, and 32 human faces (Figure 1). 22 subjects participated in the MEG experiment. Stimuli were presented for 200ms with a variable ISI of 1-1.5s. In order to maintain attention, on each trial, the stimulus was rotated slightly to the left or right (±3 deg) and subjects reported the direction of tilt with a keypress. Each of the 96 stimuli was repeated 4 times per run in pseudorandom order, for a total of 384 trials per run. Subjects completed 6 runs in total, producing 24 trials per exemplar. MEG preprocessing and analysis was conducted in MATLAB with CoSMoMVPA (Oosterhof et al., 2016).

For comparison with the MEG data, the stimuli were passed through visual saliency models (GBVS and Itti-Koch-Niebur) and the pre-trained DNNs AlexNet (Krizhevsky et al., 2012), VGG-19 (Simonyan et al., 2014), and GoogleNet (Szegedy et al., 2015) using the MATLAB implementations from Harel et al. (2006) and MatConvNet (Vedaldi et al., 2015) respectively. We used representational similarity analysis (Kriegeskorte et al., 2013) to examine the temporal dynamics of the neural response to the stimuli and to compare this to the output from the computational models.

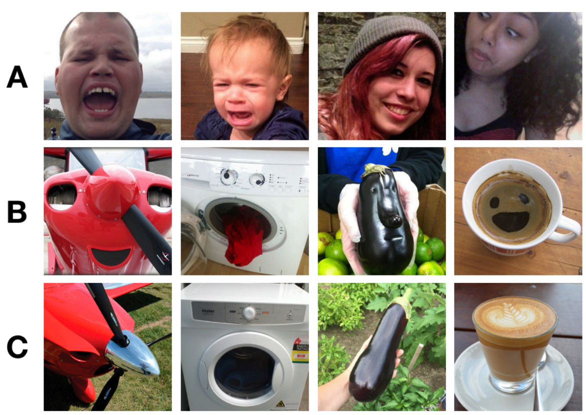

Figure 1: Example visual stimuli: (A) human faces, (B) illusory faces in everyday objects, (C) matched objects without illusory faces.

Results & Conclusions

Multivariate pattern analysis was applied to the whole-brain MEG response across all 160 sensors. Classification performance peaked at ~160-170ms post-stimulus onset for distinguishing both human faces (Figure 1A) and illusory faces (Figure 1B) from ordinary objects (Figure 1C), and for distinguishing human faces from illusory faces. At ~260ms a second peak in decoding performance occurred for discriminating human faces from objects both with and without illusory faces.

Representational similarity analysis revealed an early differentiation of human faces versus non-face objects in the activation patterns across all MEG sensors. By ~150ms after stimulus onset, some illusory face exemplars elicited MEG activation patterns that were more similar to non-face objects and others were more similar to human faces. However, this representation quickly evolved, and less than ~100ms later, all objects were represented similarly regardless of whether they contained an illusory face or not.

Both models of visual saliency (GBVS, Itti-Koch-Niebur) showed an early and transient correlation with the MEG data, with a sharp peak at ~100ms post stimulus onset. In contrast, the DNNs (AlexNet, VGG-19, GoogleNet) had a much greater and sustained correlation with the MEG data across the time-course, which generally increased for deeper layers. However, simple categorical models which grouped the stimuli into faces and objects showed a stronger correlation with the MEG data than even the highest DNN layers. This was the case regardless of whether the illusory face images were grouped with the real faces in the categorical model, or with the matched objects.

Together the results demonstrate that illusory faces are rapidly processed by the human brain, however, the underlying neural representation is only partially captured by existing computational models.

Acknowledgements

This work was supported by the Intramural Research Program of the National Institute of Mental Health (ZIA-MH-002909), the ARC Center of Excellence for Cognition and its Disorders, the KIT-Macquarie Brain Research (MEG) Laboratory and an Australian NHMRC Early Career Fellowship awarded to S.G.W.

References

Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. *NIPS Proceedings, 19*.

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*, 401–412.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012).

Imagenet classification with deep convolutional neural networks. *NIPS Proceedings, 25*.

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*:27.

Simonyan, K., & Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv*, *1409.1556.*

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). Going deeper with convolutions. *IEEE Computer Vision and Pattern Recognition, 7,* 1-9.

Taubert, J., Wardle, S.G., Flessert, M., Leopold, D., & Ungerleider, L.G. (2017). Face pareidolia in the rhesus monkey. *Current Biology*, *27*, 2505-2509.

Vedaldi, A. & Lenc, K. (2015) MatConvNet - Convolutional Neural Networks for MATLAB. *Proc. of the ACM Int. Conf. on Multimedia*.

Wardle, S.G., Seymour, K. & Taubert, J. (2017). Characterizing the response to face pareidolia in human category-selective visual cortex. *bioRxiv. https://doi.org/10.1101/233387*