

Combining Generative Adversarial Networks (GANs) with behavior and brain recordings to study scene understanding

Symposium Presentation: Friday, May 17, 2024, 5:00 – 7:00 pm, Talk Room 2

Session: [Using deep networks to re-imagine object-based attention and perception](#)

[Schedule of Events](#) | [Search Abstracts](#) | [Symposia](#) | [Talk Sessions](#) | [Poster Sessions](#)

Melissa Le-Hoa Võ¹, Aylin Kallmayer¹; ¹Goethe University Frankfurt

Our visual world is a complex conglomeration of objects that adhere to semantic and syntactic regularities, a.k.a. scene grammar according to which scenes can be decomposed into phrases – i.e, smaller clusters of objects forming conceptual units – which again contain so-called anchor objects. These usually large and stationary objects further anchor predictions regarding the identity and location of most other smaller objects within the same phrase and play a key role in guiding attention and boosting perception during real-world search. They therefore provide an important organizing principle for structuring real-world scenes. Generative adversarial networks (GANs) trained on images of real-world scenes learn the scenes' latent grammar to then synthesize images that mimic images of real-world scenes increasingly well. Therefore GANs can be used to study the hidden representations underlying object-based perception serving as testbeds to investigate the role that anchor objects play in both the generation and understanding of scenes. We will present some recent work in which we presented participants with real and generated images recording both behavior and brain responses. Modelling behavioral responses from a range of computer vision models we found that mostly high-level visual features and the strength of anchor information predicted human scene understanding of generated scenes. Using EEG to investigate the temporal dynamics of these processes revealed initial processing of anchor information which generalized to subsequent processing of the scene's authenticity. These new findings imply that anchors pave the way to scene understanding and that models predicting real-world attention and perception should become more object-centric.