



Department of Psychology  
The University of Hong Kong

香港大學心理學系

## Brown Bag Lunchtime Seminar (Theme: Cognition and Neuroscience)

### Human Cognition Driven AI Interpretability Study

12:30 p.m. – 1:30 p.m. | November 11, 2022 (Friday)  
Rm 813, 8/F, The Jockey Club Tower | Centennial Campus | The University of Hong Kong



**Yumeng Alice YANG**

MPhil student

Department of Psychology

The University of Hong Kong

#### Abstract

AI models are involved in many modules in the autonomous driving system, but the black-box nature of deep learning and AI sets back the landing of AI technologies and the trust among users. The study compares AI and humans in performance and attention strategies under driving scenarios to help AI develop better methods and enable XAI (Explainable AI) to provide better explanations. We asked 30 participants, with their eye movements recorded, to detect all the vehicles in 160 images, rated on three domains, including complexity, occlusion, and degradation, from the BDD100K dataset. We trained AI model Yolov3 on the same dataset and compared human and AI performance (the number of correct detections divided by the total number of vehicles in the image). We also compared human attention maps to AI's saliency maps. We found that AI is generally better than humans at detecting vehicles under complex, occluded, and degraded conditions. While AI is not affected by these conditions, humans are. Interestingly, high similarity to human attention strategy in degraded conditions (i.e., blurry) is associated with better AI performance.

#### About the speaker

Alice Yang is a first-year MPhil student under the supervision of Dr. Janet H. Hsiao.

#### Zoom (For participants who couldn't attend the Seminar in person)

<https://hku.zoom.us/j/3951550048?pwd=SncvL3RYakEycUtpL29vdDJEclEwdz09>

Meeting ID: 395 155 0048 | Password: psyc



~All are Welcome~

Enquiry: [rpsyc@hku.hk](mailto:rpsyc@hku.hk)