ARIA

APPLIED RESEARCH IN ACTION

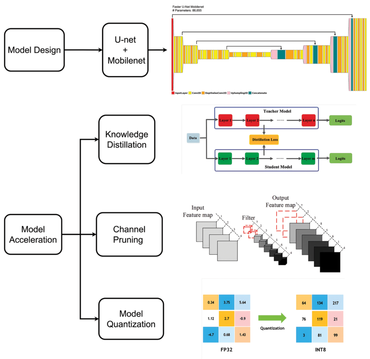
AI-Based Saliency Prediction for Better Perceptual Quality in Display and Video Compression

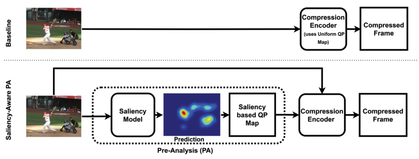
Ultimate visual experience, from natural scenes to gaming contents and beyond, with our real-time saliency prediction-based quality improvement.

Adeem Jassani

Qiang Sun
ACADEMIC SUPERVISOR

Tim Chen
INDUSTRY SUPERVISOR





PROJECT SUMMARY

Visual saliency prediction is an active research area with applications in selective quality improvement features, such as image and video compression, and backlight local dimming. Despite its importance, current saliency prediction datasets predominantly consist of real-world scene images, with limited video samples and no game-scene content. This limitation is due to the challenges of human annotation and the need for specialized hardware such as eye trackers. One of the aims of the project is to address this gap by creating a gaming-specific dataset and developing models optimized for limited gaming data to enhance saliency prediction performance.

Another critical aspect of saliency prediction is the need for low-latency models, given the real-time requirements of applications like video compression pipelines. To this end, we are optimizing the latency of existing image saliency models trained on publicly available natural scene data [1]. We have achieved approximately a 10% improvement in loss through knowledge distillation and up to a 3.8x flops speedup with an 11% deterioration in loss via channel pruning [2] for convolutional neural networks.

Building on the earlier point about data scarcity, the field faces a significant challenge in the form of limited large datasets; even the largest image saliency datasets [1] contain only 10,000 training and 5,000 validation data points. To mitigate this, we are exploring methods to generate synthetic data in addition to creating new datasets.

REFERENCES

[1] M. Jiang, S. Huang, J. Duan, and Q. Zhao, "SALICON: Saliency in context," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 1072–1080.

[2] Zhuang Liu, Jianguo Li, Zhiqiang Shen, Gao Huang, Shoumeng Yan, and Changshui Zhang. Learning efficient convolutional networks through network slimming. In Proceedings of the IEEE international conference on computer vision, pages 2736–2744, 2017.

AMD

Computer Science
UNIVERSITY OF TORONTO

Master of Science in Applied Computing