ARIA
APPLIED RESEARCH IN ACTION

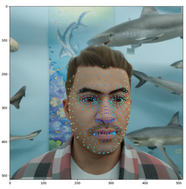
Real-Time 3D Landmark Regression from Monocular Image Input on Mobile GPUs with Synthetic 3D Face Data

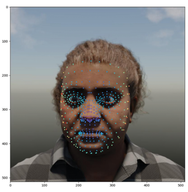
A work that infers positions of 3D face landmarks from one 2D image of the human subject in real-time

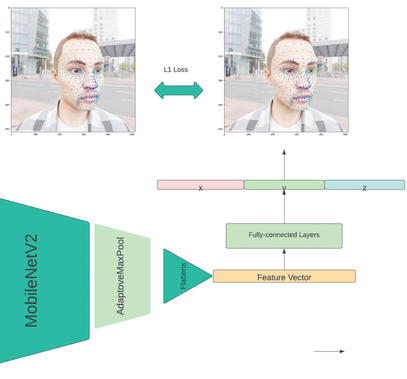
Yuqing (Sophia) Zeng

Vicky Yu and Irene Jiang
ACADEMIC SUPERVISORS

Paul Gries
INDUSTRY SUPERVISOR







PROJECT SUMMARY

Describing human face shapes and positions from face images remains a popular topic at the intersection of computer vision and deep learning. While 2D face alignment and facial landmark localization has become a well-studied field, 3D face alignment and landmark regression stays a challenging domain due to the difficulty estimating a face's 3D position and surface topology through a projected perspective. Existing approaches are restrained by the limited in-the-wild face data with manually-annotated 3D landmarks.

This project devised a light-weight end-to-end model that takes an image and output 3D landmark coordinates. The model utilized MobilenetV2 as the image-encoder backbone and is trained on a fully synthetic face dataset with dense 3D annotations and evaluated on the test split of the same dataset. Visual demos suggest the trained model is generally capable of describing the position and shape of a 3D face object through 2D projections. We also benchmarked the model on AFLW2000-3D dataset, a popular benchmark in 3D facial landmarks and alignment and achieved comparable results with SOTA methods.

MODIFACE

Computer Science
UNIVERSITY OF TORONTO
Master of Science in Applied Computing