Enabling Computer Vision Driven Assistive Devices for the Visually Impaired via Micro-architecture Design Exploration

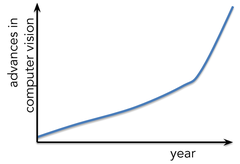
Linda Wang and Alexander Wong {ly8wang, a28wong}@uwaterloo.ca
University of Waterloo, Waterloo Artificial Intelligence Institute

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

Introduction

Background

- Rapid improvements in computer vision
- Potential to aid in tasks where previous solutions struggled → Assistive devices for visually impaired
- Challenges: large size, privacy, affordability, widespread deployment



Goal

A network architecture exploration approach is investigated to create an optimal object detection network for OLIV

- Numerical optimization problem: find a set of hyperparameters that optimizes the trade-offs between accuracy, speed, and size of the deep neural network.

Micro-architecture design exploration

- Exploration of hyperparameters, width and resolution multiplier, to find a MobileNetV2-SSD network that offers the best trade-offs between performance, speed and size

$$\underline{\theta} = \{\alpha, \rho\}$$

width multiplier, resolution multiplier

- To address the trade-off between accuracy, speed and size, the micro-architecture exploration approach is formulated as a numerical optimization problem

$$\Omega(\mathcal{N}(\underline{\theta})) = 20 \log \left(\frac{a(\mathcal{N}(\underline{\theta}))^{\kappa}}{p(\mathcal{N}(\underline{\theta}))^{\beta} r(\mathcal{N}(\underline{\theta}))^{\gamma}} \right)$$

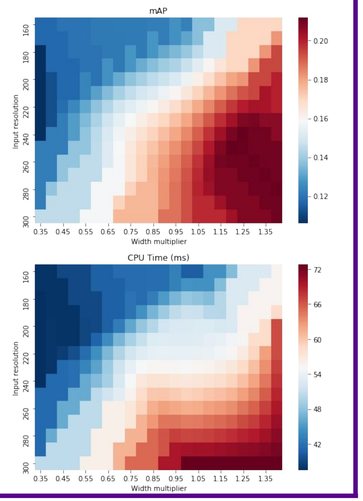
modified NetScore, accuracy of network, # of params in network, CPU running time

- The formal objective function used to determine the most optimal hyperparameters for this application is defined as:

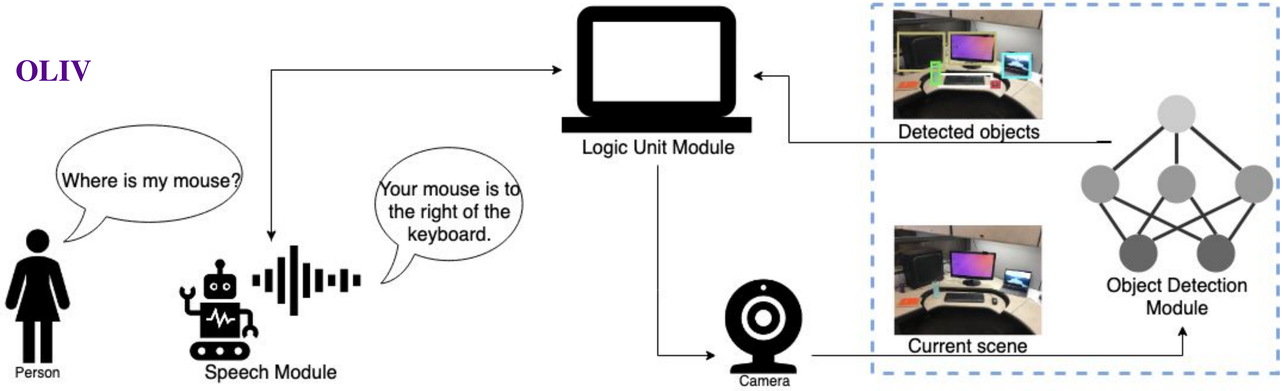
$$\hat{\underline{\theta}} = argmax_{\underline{\theta}} \Omega(\mathcal{N}(\underline{\theta}))$$

Width and Resolution Multiplier

- As the width multiplier and input resolution increase, the mAP and CPU run-time also increase
- mAP plot: there is little increase in mAP after the width multiplier reaches 1.15 and input resolution reaches 220
- CPU time plot: run-time continues to increase linearly as the width multiplier and input resolution increase

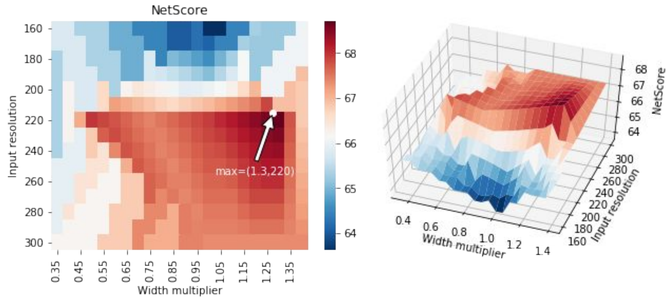


OLIV



Results and Discussion

- Width multiplier of 1.3 and input resolution of 224 achieved the highest score → balanced trade-offs for OLIV
- Resulted in a well-suited compact object detection model that offers a balanced trade-off between accuracy, speed and size
- Potential to be applied to other assistive devices to offer users a cost-efficient and secure solution



This work was supported by the Natural Sciences and Engineering Research Council of Canada and Canada Research Chairs Program. The authors would also like to thank Microsoft.