

Deep Learning to identify and Predict Object in the Environment

Nicholas Sagers¹, Juan Andrade², Robert LiKamWa^{1,2}, and Andreas Spanias²

¹School of Arts, Media and Engineering (AME), Arizona State University

²SenSIP Center, School of ECEE, Arizona State University

Abstract—The intersection of deep learning and mobile/embedded systems has received significant attention in recent years. Development in the areas of deep learning and mobile architecture has made possible an entirely new ecosystem of technological solutions. We propose one of those solutions, a mobile system to predict the whereabouts of objects in the environment. Using recently developed techniques, we are able to efficiently utilize computer vision and mobile processing to assist in the task of interacting with the natural environment. This solution provides a means of assisting those with memory impairments and those without memory impairment alike.

Index Terms—accessibility, augmented reality, computer vision, mobile, convolutional neural networks

I. INTRODUCTION

New object detection and natural language processing techniques have allowed for a plethora of new, powerful, and progressive techniques and technologies, such as energy requirement reduction via selective image processing. [1, 2]. These two innovations are foundational for future applications that involve more naturalistic human-machine interfacing. Existing technologies have utilized object detection for security, automation, and entertainment and natural language processing for hands-free interaction and real-time translation. We propose a system to improve the quality of life of both memory impaired and non-memory impaired persons through the use of deep learning techniques and mobile systems.

II. IMPLEMENTATION

The Real-World Search Engine (RWSE) utilizes the Microsoft HoloLens in order to provide visual feedback to the user and as a source of data for a network connected neural network. The HoloLens is a relatively recent hardware development that has numerous, unexplored uses in industry and accessibility [3]. Techniques such as down-sampling [4] and motion estimation will be used to ensure a positive user experience in terms of energy performance and functionality. Object identification and natural language processing will be carried out via already-existing models such as Squeezenet [5]. This project will focus on the application of these models, the feasibility and implementation of deep learning tasks on low powered devices (Microsoft HoloLens), and future applications of this system. The future challenges of the project, outside of design and testing, is the meeting of the power and bandwidth requirements.

ACKNOWLEDGMENT

This program is funded by NSF CISE award 1659871



Figure 1. Prototype example view of Real World Search Engine.

REFERENCES

- [1] M. Buckler, P. Bedoukian, S. Jayasuriya, and A. Sampson, "EVA²: Exploiting temporal redundancy in live computer vision," in *2018 ACM/IEEE 45th Annual International Symposium on Computer Architecture (ISCA)*, 2018, pp. 533–546.
- [2] Y. Zhu, A. Samajdar, M. Mattina, and P. Whatmough, "Euphrates: Algorithm-soc co-design for low-power mobile continuous vision," *preprint arXiv:1803.11232*, 2018.
- [3] G. Evans, J. Miller, M. I. Pena, A. MacAllister, and E. Winer, "Evaluating the microsoft hololens through an augmented reality assembly application," in *Degraded Environments: Sensing, Processing, and Display*, vol. 10197. International Society for Optics and Photonics, 2017, p. 101970V.
- [4] X. Chen and A. Gupta, "An implementation of faster RCNN with study for region sampling," *preprint arXiv:1702.02138*, 2017.
- [5] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size," *preprint arXiv:1602.07360*, 2016.