**Project Proposal**

**1.** **Project Title:**

Intelligent Checkout System

**2.** **Team members (names and emails):**

James Bozhkov - zbozhkov@students.kennesaw.edu

Jiaming Li - jli36@students.kennesaw.edu

**3.** **What is the problem/question that you will be investigating?**

We will be investigating how to use object detection and face recognition techniques for the implementation of a smart self-checkout system that has the ability to improve checkout efficiency. Specifically, two problems are needed to be solved. First and foremost, how to use objection detection techniques to precisely examine the merchandise customers purchase. Secondly, how to extract, retrieve and compare facial features of people to determine the identity of customers.

**4.** **What are the most relevant readings (initial references)? (5-8 papers)**

[1] E. Winarno, W. Hadikurniawati, I. H. Al Amin, and M. Sukur, “Anti-cheating presence system based on 3WPCA-dual vision face recognition,” in 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Sep. 2017, pp. 1–5, doi: 10.1109/EECSI.2017.8239115.

[2] M. Y. F. Siddiqui and Sukesha, “Face recognition using original and symmetrical face images,” in 2015 1st International Conference on Next Generation Computing Technologies (NGCT), Sep. 2015, pp. 898–902, doi: 10.1109/NGCT.2015.7375249.

[3] C. Liu, “The development trend of evaluating face-recognition technology,” in 2014 International Conference on Mechatronics and Control (ICMC), Jul. 2014, pp. 1540–1544, doi: 10.1109/ICMC.2014.7231817.

[4] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” ArXiv180402767 Cs, Apr. 2018, Accessed: Sep. 01, 2020. [Online]. Available: http://arxiv.org/abs/1804.02767.

[5] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” ArXiv150601497 Cs, Jan. 2016, Accessed: Sep. 01, 2020. [Online]. Available: http://arxiv.org/abs/1506.01497.

[6] H. Liau, N. Yamini, and Y. Wong, “Fire SSD: Wide Fire Modules based Single Shot Detector on Edge Device,” ArXiv180605363 Cs, Dec. 2018, Accessed: Sep. 01, 2020. [Online]. Available: http://arxiv.org/abs/1806.05363.

**5.** **What data will you use?**

To train the objection detection model, we are going to use MVTec Densely Segmented Supermarket Dataset (MVTec D2S). D2S contains 21,000 high-resolution images that belong to 60 categories. The objects include groceries and everyday products which are exactly what is necessary for this project. The dataset includes images with pixel-wise labels of all object instances which can be used to train the objection recognition model.

**6.** **What are the existing methods? Are their implementations available?**

The barcode based self-checkout system has been widely utilized in grocery stores including Walmart, Target and Kroger. The machine vision based system has a potential to achieve higher efficiency but it has not yet been fully adopted. The software company [Mashgin](https://www.mashgin.com/) offers a Touchless Checkout System solution using AI. They advertise that their system can identify any items without the need of a barcode, and are 10 times faster than the traditional checkout method. Since it is an industry software, we don’t have access to their implementation. We assume that they are using the object detection model for identifying items by applying face recognition techniques for determining each customer’s identity. Therefore, we are planning to implement this system based on our assumptions.

**7.** **What method or algorithm will you use?**

There are a lot of state-of-the-art deep learning based object detection models which we can use. For example, You Only Look Once (YOLO), R-CNN and Single Shot MultiBox Detector (SSD). We will evaluate each one of them and integrate the one with the best optimal performance into our project.

For the face recognition part, we can use the CNN based face detection and facial landmark detection model provided by dlib. But we are open to investigate the latest research for accomplishing the burden of face recognition.

**8.** **How will you evaluate your results?**

We will first evaluate the accuracy of the object detection model. More specifically, we will use the model for predicting the type and count of items in the input image. The items will be placed overlapping to increase the prediction difficulty in order to ensure the reliability of our model. Similarly for the face recognition model, we will first import several people’s face features into our system. Then we will test if the system can determine the correct identity. The faces that are not embedded in the system will also be tested to confuse the system. In addition, the processing time is also a key factor we need to evaluate. We will compute the average processing time cost of each stage in our system. For example, the processing time of object detection and the time cost for determining a customer’s identity.

**9.** **What kind of analysis will you use to evaluate and/or compare your results (e.g. what performance metrics or statistical tests)?**

Average Precision (AP) is a metric which is used to evaluate the performance of object detection models. It computes the average precision value for recall value from 0 to 1. In our scenario, it represents how well the model predicts the correct label. We will select a high AP due to it being very critical for calculating a correct subtotal.

Recognition Rate can be used for evaluating the face recognition algorithm. Recognition Rate is the total number of correctly identified probe images divided by the total number of probe images.