Computer Vision Deliverable 2

Git: https://gitlab.com/jagadeeshsudhakar/analysing_shopping_pattern

Problem Statement: Our groups focus will be analysing customer shopping patterns to increase time spent in stores.

2.1 Research

Your research should be spread across a few key areas. Depending on your project, your research may be more "business focused" or more "academically" focused. You should address at least 10 of the questions below, and must choose and review 3 academic papers.

2.1.1 Business/Customers

- What problem will your Computer Vision solution solve, and for whom?
    - We try to analyze customer shopping behavior by monitoring their activity using CCTV footage. Algorithms developed would be focused mostly on offline learning where data gathered would be analyzed later for insights.
- What value will it provide them? What are their pain points?
    - Value: Outcomes of the algorithm should help the store better understand the shopping behavior of customers and therefore help in stocking products that would maximize their chances for a higher profit.
    - Pain: Similar to most computer vision projects, privacy can be an issue. We may have to think of ways to keep fingerprinting of users to a minimum. In addition to this, integrating this data into the already existing CCTV system must be seamless.
- How big is the potential market?
    - The market for this sort of technology includes all brick-and-mortar stores, warehouses, and factories. If people walk around in an area, understanding the patterns of this movement can lead to optimization, whether it is aimed to increase productivity in a warehouse or factory, or making more sales in a department store.
    - In addition, our solution can be implemented with security cameras, an infrastructure component that exists in the large majority of these settings. This easy barrier-of-entry to the technology may be appealing to potential customers.
- Do other similar solutions exist?
    - There are a few different solutions in this arena, but none have been able to get a good foothold quite yet. One such solution is provided by Walkbase, a company

that uses WiFi and bluetooth to triangulate customers in the store via their cell phones.
    - Some stores use different techniques to make their customers stay and shop more. Some of these solutions are implementing a restaurant or lunch place inside the store. Another idea is to have a digital section with a screen where customers can access the store website and purchase items via the website.(https://www.vendhq.com/blog/get-customers-to-spend-more-time-and-money/)
- Would your business have any competitors? Who are they? How are they doing?
    - One excellent solution is Amazon Go which tracks customers from the start to finish by fingerprinting the customers across aisles to implement an instant shopping experience where customers can just walk out of the store and get billed automatically for whatever they shopped. (https://www.amazon.com/b?ie=UTF8&node=16008589011)
    - Another closely competitive solution is by Philips and by interact-lighting which implements an Indoor GPS like positioning systems respectively. These systems have the highest accuracy of indoor positioning among other solutions with the accuracy of almost 30cm with which the solutions can track the users with a great precision and hence helping the supermarket for shopping patterns.
    - (https://www.usa.lighting.philips.com/systems/lighting-systems/indoor-positioning and https://www.interact-lighting.com/global/what-is-possible/interact-retail/indoor-navigation)
- How are potential customers dealing with these issues now?
    - For businesses that decide consumer tracking is useful, many solutions that are provided either end with analysis of sales or implement wifi and bluetooth triangulation to locate customers over the period of time that they are in the store. The former solution misses a large amount of data, and the latter solution requires expensive infrastructure that reduces the viability of the product for many potential customers.
- How much would a customer be willing to pay for your product?
    - A business would be willing to pay for our product as long as it costs less than the economic benefits that they will receive for implementing such a system. Due to not needing new hardware, the entry cost for our product would be quite low. As a business strategy, it would be possible to provide our product to a store as a trial. When the trial is up, the business would be able to decide if they would like to pay and continue using our product.

- Are your customers individuals or businesses? Be sure to cite the sources you reference in your research.
    - Our customers are businesses because our solutions allow them to understand trends in their customers and allow them to more effectively market toward them. This is self-evident in what our product provides, as only businesses have stores.

2.1.2 Academic Literature Review

A thorough literature review can save weeks of wasted time implementing dead-end solutions or re-doing work that others have already done. You should pick 3 academic papers (we can help you with this, if needed), and do a "deep-dive" on these publications.

- What academic work is relevant to your project topic? Pick 3 papers, ask us for help if you need it.
    - Methods for Augmenting Transaction Data with Visually Extracted Demographics of People, Using Computer Vision. (*Authors*: Rajeev Sharma, Hankyu Moon, Varij Saurabh, Namsoon Jung, et al.) [https://patents.google.com/patent/US8010402B1/en]
    - A Vision System for Automated Customer Tracking for Marketing Analysis: Low Level Feature Extraction (*Authors*: Alex Leykin, Mihran Tuceryan) [http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.65.6454&rep=rep1&type=pdf]
    - Shop and Go: An Innovative Approach Towards Shopping Using Deep Learning and Computer Vision (*Authors:* Nerenda Shekokar, Anchal Kasat, Sanjay Jain, Pranjal Naringerekar, et al.) [https://ieeexplore.ieee.org/abstract/document/9214256]
    - Singh, B., Marks, T.K., Jones, M.J, , Tuzel, C.O.,"A Multi-Stream Bi-Directional Recurrent Neural Network for Fine-Grained Action Detection",IEEE Conference on Computer Vision and Pattern Recognition (CVPR),DOI:10.1109/CVPR.2016.216,June 2016,pp. 1961-1970. [https://www.merl.com/publications/docs/TR2016-080.pdf]
- What makes these papers important/relevant?
    - The research gathered juxtaposes historical computer vision methodologies with more current practices. Researching the evolution of these practices enables us to recognize and adopt computer vision applications consistent with our groups solution statement. It gives us a better understanding of relevant and more effective feature extraction techniques to-date. With that said, these research papers are not absent of any actual applicable computer vision techniques by any

means. In fact, these sponsored research projects relevant to our research, written almost two decades ago, are still active.

- What are their results and how did they achieve these results?
    - The researchers were able to collect demographic data at the point of sale utilizing cameras that would determine the demographic class of a particular person. The algorithm would determine group information--buyer versus bystander--and extract facial feature data to determine a person's overall emotional response during the transaction. They achieved this by measuring visually perceptible attributes that were within range of their sensors.
- What's different/unique about these approaches?
    - These papers aren't necessarily different or unique, they are simply research into the different technical challenges that we will face along the way. They are simply acting as a knowledge base as we build our product.

2.1.3 Open Source

- What open source code is available that is relevant to your topic?
    - There is a large codebase for human identification, especially for PyTorch, found here- https://github.com/layumi/Person_reID_baseline_pytorch
- How active are the communities around this code?
    - Not particularly active. There seem to be small fixes every so often, but there don't seem to be many contributors. This may be due to lack of training data.
- What data is available for testing and/or training algorithms?
    - Raw Data - https://www.merl.com/demos/merl-shopping-dataset
- Is labeled data available? How much? How is the data licensed? Is it under copyright protection?
    - There are currently no labels available for the dataset
    - The MERL Shopping Dataset contains 106 videos, each of which is a sequence ~2 minutes long. 41 people have participated in the shopping videos, with SubjectID = 1,...,41. Each subject is in up to 3 videos, with SessionID = 1, 2, 3. Each of the 106 video files has a filename of the form `xx_yy_crop.mp4`, where xx is the SubjectID and yy is the sessionID.
    - The data has to be labelled manually by reading each frame from the video to determine the actions.

Permission to use these data and their documentation without fee for research and educational purposes, is hereby granted.

In no event shall merl be liable to any party for direct, indirect, special, incidental, or consequential damages, including lost profits, arising out of the use of this dataset and its documentation, even if merl has been advised of the possibility of such damages.

Merl specifically disclaims any warranties, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. the data provided hereunder are on an "as is" basis, and merl has no obligations to provide maintenance, support, updates, enhancements, or modifications.

2.1.4 Industry Solutions

- What companies are solving similar problems to yours?
    - Mitsubishi, AmazonGo, Philips, Interact-Lighting, Walkbase.
- It can be tough to tell exactly how proprietary solutions work, but what can you find on the internet?
    - Tracking people from video footage isn't cutting edge technology, but combining this with customer analytics is. There is little data on this on the internet.
- Has anyone reverse engineered these products?
    - Not that we know of. This is quite a specialty field, which may be a limiting factor.
- If you have access to the product, what can you learn from using it?
    - Not Applicable.
- Are there available talks, documentation, or other resources from their engineering teams?
    - Not publically.

Data Preprocessing:

We have found a dataset which is explained above but it has to be manually labelled since the previous work of that dataset was done with Matlab and there were no available data points to continue from. Manual labeling is in process. Since there are 106 videos it takes time to label data points. We should be able to complete this along with the next deliverable.
We have labelled one such video to show how the final labelling would look like.