Identifying characteristics of passer-by's to provide dynamic advertising in public spaces using Computer Vision

UNIVERSITY OF LINCOLN

Introduction

Online advertising has seen exponential increases in effectiveness over recent years, as a result of advertising services (spearheaded by Google) creating profiles of each individual's online identity, such as their basic details (gender, age etc) and their interests, for example if they watch sports or search for cooking recipes. Advertisers are then encouraged to use this information to scope into their specific demographic, such as beer companies advertising to young American males interested in sports around the time of major sporting events such as the Super Bowl. This means the advertiser only has to pay for those who see the ad, instead of rolling it out across the entire site and having it only apply to a small number of the people viewing it, despite paying for full site coverage. Advertising on digital screens in public spaces is still vulnerable to these issues, as these devices have no means of discerning the characteristics of passers-by, and therefore wouldn't be able to change the advertisement as a result. The primary issue is that, unless the user can be persuaded to stop and sign into an online account (such as social media), there is no easy way of collecting the users' data. Therefore, more complicated means are required, such as using Bluetooth to scrape data from the phone [1], or using Computer Vision to analyse the passer-by. This can't provide the wealth of data that online advertising can, however it is not unreasonable to expect it to be able to build a basic profile, such as age, gender and facial features. This minimal amount of information would allow ads to be tailored to the profile of the person in front of them, meaning adverts could be significantly more cost effective.

The problem is made more complex by the variety of different types of faces, and sizes of groups

1: https://www.popsci.com/article/technology/hundreds-bluetooth-beacons-secretly-track-new-york-city-passersby

Aims and Objectives

The primary aim of this project is to develop a system which integrates tailored advertisements with digital signage to increase the return on investment of ads in public spaces.

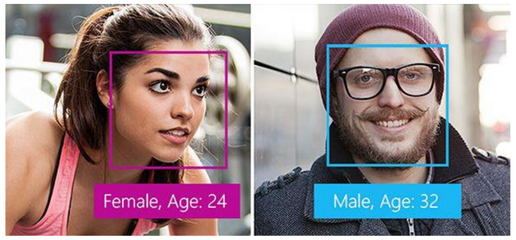
To achieve this aim, the following objectives have been laid out:

Objective 1: Develop a system which can find faces on a photo.
Objective 2: Identify the individuals' features, such as age, gender and hair.
Objective 3: Gather secondary data, such as location and time
Objective 4: Infer the context of groups, such as family, friends or couple.
Objective 5: Use the gathered information to recommend an advert
Objective 6: Run this system in real time

The aim would be measured by placing the system in a physical setting, and tracking the amount of time spent looking at the ad.

Research Methods

This project will largely focus on the identification of facial characteristics, in order to predict the demographics of that passer-by. Therefore, different computer vision techniques and API's will be identified and compared to see how they perform in a public setting. Examples of this are Google Vision, Microsoft Azure Face API and Amazon Rekognition, as well as the techniques behind them, which is usually a combination of computer vision to extract features, and machine learning to identify demographics based on this. From here, the project will try to recommend a suitable advert for the passer-by, and the results of this will be measured for accuracy by being tested against humans, who should have similar levels of accuracy. Complications may include noise/obscured faces, as well as other extraneous variables associated with running a system in public, such as weather and damage.



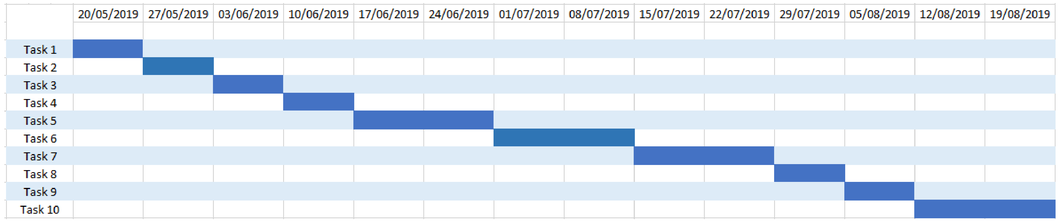
An output of the Microsoft Azure Vision API

Research Plan

To achieve the aim and objectives, the following steps have been laid out, that need to be completed as part of the project:

1. Identify faces in an image
2. Get features of the face
3. Get other features, such as hair
4. Retrieve information, such as time and location
5. Use information to recommend an appropriate advertisement
6. Evaluate performance compared to people
7. Infer context of the group (e.g. family, friends, couple etc)
8. Use this information to make more accurate recommendations
9. Compare the old recommendations to the new
10. Implement the system in a real time, public environment

To plan the timespan of the project, a Gantt chart was creating based on these tasks, as shown below. This identifies 14 weeks to complete the project, with the first four tasks taking up the first month (as this will require the implementation of multiple API's and methods), and a larger time dedicated to tasks that won't have a pre-existing solution, such as making an advertisement recommendation, or inferring the context of a group of people, based on observed information.



A Gantt chart showing the predicted progress of the systems development