Extracting Key Demographics Using Computer Vision

Neil Mayo, Frankie Inguanez
Institute of Information & Communication Technology
Malta College or Arts, Science & Technology
Corradino Hill
Paola PLA 9032
{neil.mayo, frankie.inguanez}@mcast.edu.mt

***Abstract*—In this study, we propose that key demographics such as age and gender can be extracted for implicit and explicit data gathering. During this study, it was discovered that this is possible however it is not always accurate due to other factors. A prototype was created using python and OpenCV which can be fed live footage or images of a face and the age and gender related to that face would be displayed using a CNN network architecture and the Adience dataset used to train the model. In the next stage of this study more demographics will be detected, this will be done to not only be able to gather more data on an individual but to also help find which other demographics and factors are problematic.**

***Index Terms*—CNN, Python, Adience, Age detection, Gender detection, Facial recognition, Data Gathering**

I. Introduction

Computer Vision has become common as facial detection software being used with the recent progress in facial recognition software integrated into phones has become normal and with stores such as amazon go introduced, a fully automated grocery store allowing customers to walk in and walk out with their purchases meaning there are no queues or employees such as cashiers [1]. Thus, we are posing the following hypothesis in this study: It is being assumed that key demographics can be extracted using computer vision, Although there are multiple different demographics the focus will be on age and gender, accuracy and issues encountered will be investigated for the extraction of key demographics to be integrated into stores and other buildings such as workplaces for explicit and implicit data gathering. To perform this study, three primary questions must be addressed:

1) What features in a data set determine an accurate detection of demographics?
2) which demographics are problematic or too similar to detect automatically?

In Section II, the extraction of age and gender will be researched. Section III will explain the developed prototype, how it works and answer what features will affect the accuracy of the obtained age and gender and which demographics are problematic. The findings and observations will be documented in Section IV and finally, in Section V the final thoughts and future improvements to the prototype will be discussed.

II. Literature Review

The extraction of age and gender has been regarded as a critical module for a wide variety of computer vision applications, including monitoring and passive demographic data gathering. When the age is retrieved, rather than the exact age being returned an age group is instead, as it is harder to accurately detect the exact value of an individuals age. In the following research by Gill Levi and Tal Hassner [2], the Adience benchmark [3] was used for age and gender classification which are real images that are severely blurred or having a low resolution with different poses and expressions there are 26,580 images of 2,284 different subjects.

A Convolutional Neural Network is used for age and gender estimation having images gathered from social image repositories with access to personal information such as date of birth and gender. These images are then used to create a large image training set being labelled for age and gender estimation. The CNN architecture is a network containing rectified linear operations and pooling layers following each of three different convolutional layers with the first two layers using local response normalization.

For network training, dropout learning is applied which is when the network which also contains two dropout layers with a dropout ratio of 0.5, its neurons output value are set to zero at random. Then data augmentation is used this is done by selecting a random crop of 227 x 227 pixels from the 256 x 256 inputted image and randomly mirroring it for each forward and backward training pass and the initial learning rate of this is $e-3$, then decreased to $e-4$ after 10K iterations. Finally, two methods are used one which center crops the facial image and is fed to the network cropping 227 x 227 around the center of the face. The second method is designed to help when small misalignment's are introduced, thus improving alignment quality and this is done by extracting five 227 x 227 pixel crop areas one from the center and four from different corners of the 256 x 256 facial images along with their horizontal reflections. Afterwards, the average prediction value will be considered to be the final prediction.

III. Research Methodology

The extraction of key demographics is starting to become more common such as in security cameras and even automated

self-checkout stores such as Amazon Go. The problem is that some demographics are problematic as they are too similar or difficult to detect and some features in a data set that will determine accurate detection such as lighting, camera angle and camera distance. These demographics can be extracted such as gender and age by using video footage or images it is possible to extract a person's key demographics automatically using computer vision. The aim is to find out how accurate the extraction of key demographics in certain situations can be used for implicit and explicit data gathering as this can be used to keep track of individuals in a public location, help with tracking an individual, an individual might try and buy alcohol without identification or enter a bar. This study attempts to answer the following questions like what features in a dataset determine a proper or accurate detection of demographics and which demographics are problematic or too similar to detect automatically.

The dataset being used is an Adience dataset that was used to train the model. In this case, instead of detecting a specific age, an age group is being detected. When detecting the age and gender of an individual or multiple individuals under different circumstances the results differed.

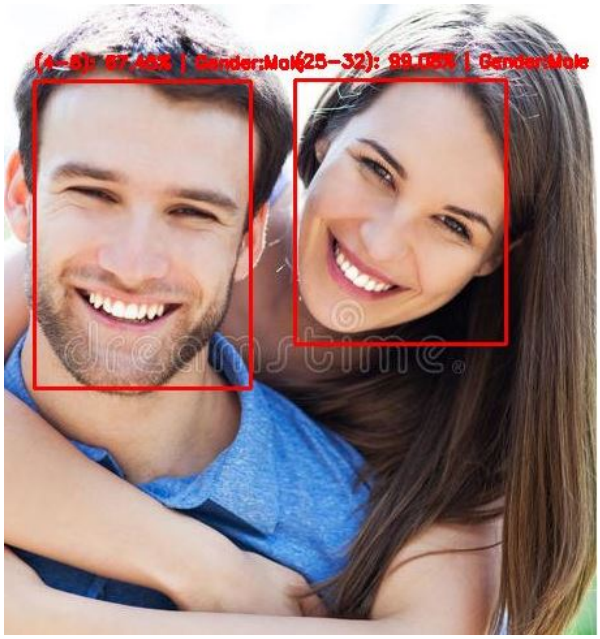

Fig. 1. Couple

When used on a couple in their late twenties, the result for the male smiling with half of his eyes closed was detected as male and aged from 4 to 5 with 57.45 percent accuracy, while the female was detected to be a female and aged from 25 to 32 with 99.08 percent accuracy, although the female's age and gender were correct the male age was wrong, in this scenario, the camera was at a good distance from the couple, the camera quality was good and the picture was taken outside under the sun resulting in good lighting even in this perfect scenario the detection was still wrong.

In a similar circumstance where there was also a couple, in the first image it resulted in both genders being correct and only the females age correct similar to the first image, however when the image was cropped to show only the males face the results were accurate, solidifying the reasoning behind there being problems in group photo.

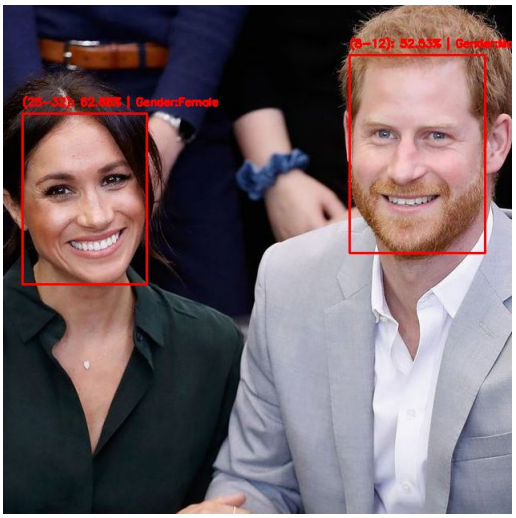

Fig. 2. Couple

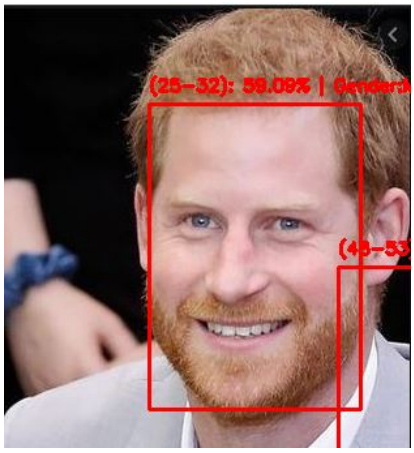

Fig. 3. Couple

In a store using CCTV footage with multiple people when trying to detect age and gender, the result is null as faces aren't being detected this is due to the distance and the camera is high up as when used on a group photo, although not fully accurate faces are being detected. In this case data gathering cannot happen as faces are not being detected.

In similar image with one person looking down and a camera facing him, his face was not only detected but also given an accurate age and gender, meaning that the resolution and upper angle is affecting the image from the store.

Another image an old man in the dark with half of his face being covered by darkness and the image being black and

Fig. 4. Store CCTV Footage

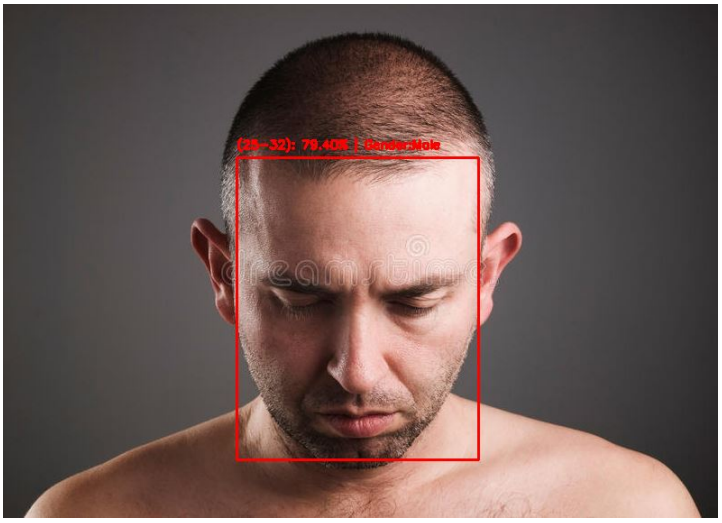

Fig. 5. Man Looking Downwards

white resulted in the age estimation being 60-100 which is accurate although the gender resulted as female.

In the last instance a man evidently wearing a wig and make up was detected as a female with their age being shown to inaccurate.

IV. Findings & Discussion of Results

The current prototype does what it is made to do, however, it is not consistent sometimes displaying correct or incorrect results. The prototype only detects two key demographics, although in the future we intend on detecting more demographics to detect, this will help us find and understand which demographics are conflicting and are problematic. Aside from conflicting demographics, there are also different factors such as lighting, distance, image quality and one which was clear was having more than two people in an image, also when used inside a small room the camera being almost height level to that person the results were consistent accurate most of the time even when the face was being covered with a hand or mask.

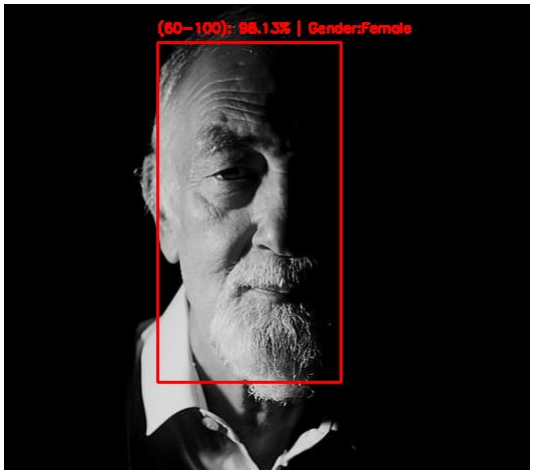

Fig. 6. Old Man In Dark

Fig. 7. Man In Makeup

V. Conclusion

We can conclude by saying that although the extraction of key demographics is possible and can be used for data gathering it is not always accurate under some circumstances. However, with a simple camera and with the right software setup you can create an age and gender detection system. The way this research can move forward and improve would be:

1) Additional key demographics to be detected,
2) Detection for features such as darkness and blur,
3) Solutions such as if image or video is dark increase brightness,
4) Data set with faces in different angles and group photos.

Therefore we can accept our initial hypothesis that the extraction of key demographics using computer vision is possible and data gathering although not always accurate can

be used and improved further. Thus, this research shall proceed further to follow up on these improvements. experience.

REFERENCES

[1] M.-P. E. M. L. R. . A. M. Fazl-Ersi, M, "Age and gender recognition using informative features of various types," p. 1, 2014. [Online]. Available: https://ieeexplore.ieee.org/document/7026190

[2] T. Levi, G Hassner, "Age and gender classification using convolutional neural networks," pp. 1–5, 2015. [Online]. Available: https://ieeexplore.ieee.org/document/7301352

[3] "Unfiltered faces for gender and age classification," 2014. [Online]. Available: https://talhassner.github.io/home/projects/Adience/Adience-data.html#agegender