Developing Testing Frameworks for AI Cameras

Arkadiusz Herdzik and Carl James-Reynolds [0000-0002-5892-5415]

1 Middlesex University, London NW4 4BT

2 Middlesex University, London NW4 4BT

C.James-Reynolds@mdx.ac.uk

Abstract. It is possible for inexpensive cameras to include AI based features such as face recognition. However, a test framework for such cameras is required that will allow comparison of accuracy under differing conditions. This will then lead to the improvement of training data and algorithms.

A simple test framework has been developed and partially evaluated by testing multiple head/face accessories under different lighting conditions. Six participants took part and 300 pictures using a Huskylens were taken under a range of conditions. It was found that the camera could detect faces at a reasonable level of accuracy during 'middle of the day' lighting conditions, with or without head accessories. However, it delivers significantly lower detection rate with accessories that cover greater parts of the face and under green light.

There is still a need to further investigate this area of study with a higher number of participants in a more controlled environment. It is anticipated that better testing frameworks will lead to better algorithms, training data and specifications for users.

Keywords: Face Recognition, Test Framework, AI Camera.

1 Introduction and Background

A test framework for face recognition technology is required in order to ensure that users can make informed choices for real life critical scenarios.

AI devices have become easily available for purchase [1] and offer a range of different possibilities including, but not limited to, automating or improving tasks conducted by humans in day to day lives [2]. Mobile phone facial biometric security and Automated Border Control Systems (also known as E-Gates) are examples of what can be achieved using algorithms and data sets for training. As another example biometric approaches for security, can be used, as each person's biometrics are unique to that specific individual and face characteristics can quickly be detected [3].

With technology such as E-Gates, it is generally acceptable for the system to be inaccurate at verifying the identity of a person, and there is intervention in the form of an immigration officer that will take control [4]. This not the case with other uses of biometrics such as those in mobile phones or door entry systems where there is no human activity that monitors what the system does and its accuracy. This raises a question

whether security systems and other AI tools are secure enough to be trusted when deployed for mass use.
There are a wide variety of strategies for implementing face recognition, which makes the need for standards in testing even more important. Some [5] use facial features, but omit hair, using luminance as the source of information; others [6] use depth as an additional variable which increases accuracy. Using a depth sensor such as the one of a Microsoft Kinect [7] is primarily needed to define the area that separates the human and the background. Researchers [8] have emphasised the importance of increasing the number of iterations for the training model to improve accuracy. Because there is such a wide range of strategies and algorithms, these are not explored, however there are many reviews of this area [9] [10] [11].

A test framework needs to consider a number of factors:

- Environmental Factors
  - Light Colour and Intensity
  - Light Direction
- Positional Factors
  - Angle and Distance of camera to face
- Human Factors
  - Ethnicity
  - Hair Styles
  - Head Coverings
  - Face Coverings
  - Make-up
  - Gender
  - Age
  - Attempts to "cheat" the system

It may be that some factors can be kept as constants, so that subjects are asked to be a specified distance from the camera and to look straight into the lens, or to remove face coverings. In real life scenarios, it is not always possible to guarantee lighting conditions and some of the "Human Factors", such as make-up, may change over time.

In order to explore the framework, a test apparatus was developed, consisting of a picture frame with a NeoPixel LED Strip [12] placed around the inside border of the frame, pointing at the subject. An Arduino Uno [13] was used to control the LEDs to allow for colour and light direction to be changed. The frame was kept one metre from the camera.

The camera used for evaluation was Huskylens [14] (Zhiwei Robotics, 2022) developed by DFRobot. It contains a 2.0 megapixel camera; it doesn't contain a depth sensor and relies primarily on an already existing hard coded algorithm. There is an option of changing the threshold for face detection, but there is no publicly available information about how this works.

Figure 1. shows the schematic of the testing system. The NeoPixel LED strip was placed inside a picture frame (facing the subject) and the process of taking the images was automated with each subject being illuminated from the top, bottom and sides using

different colours. This was simplified to red, green and blue, but the colour range could easily be extended. The subject was one metre from the camera as shown in figure 2. Two Arduino libraries were used: Huskylens [14] which is for the Huskylens camera and the Adafruit NeoPixel library was used for the addressable RGB strip. The automation of the process made testing simpler and more reliable.

For each subject 60 passport style photos were taken. During the testing, pictures of participants were stored within micro-SD card and subsequently stored securely in the cloud.

Using observational technique, each image was checked to determine whether the Huskylens detected a face within the picture, and the result was recorded.

There can be situations where there is no border at all, indicating that although there is a face in the image, the Huskylens is unable to distinguish whether it is a face or part of the background.

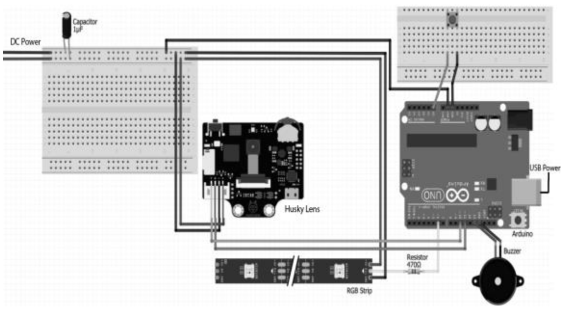

Fig. 1. Schematic of Apparatus

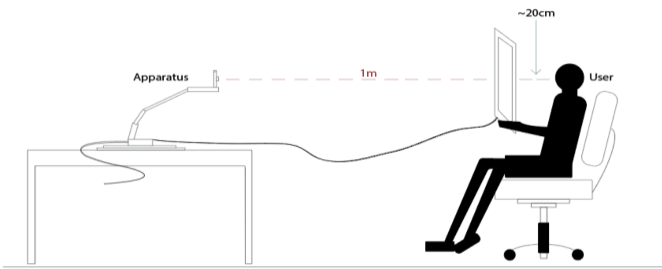

Fig. 2. Setup of Equipment

2 Test Results and Findings

There were total of 6 participants of which 3 were males and 3 females.

The main results are shown in Figure 3. It can be clearly seen that performance of the face recognition was best under red light and worst under green, which is surprising as green contributes the most to luminance. Unsurprisingly facemasks had a significant impact on performance, whilst hats had little impact, except in one case where the hat obscured the face. The camera detected a face, but could not identify it, as in figure 4.

Glasses generally reduced detection rates, but typically only by 1/6. It was noted that in tests with the female participants wearing make-up that detection rates reduced, however due to the low number of participants, it is difficult to assign any significance to this. Across all tests except those with red light and no accessories, males were twice as likely to be detected as females.

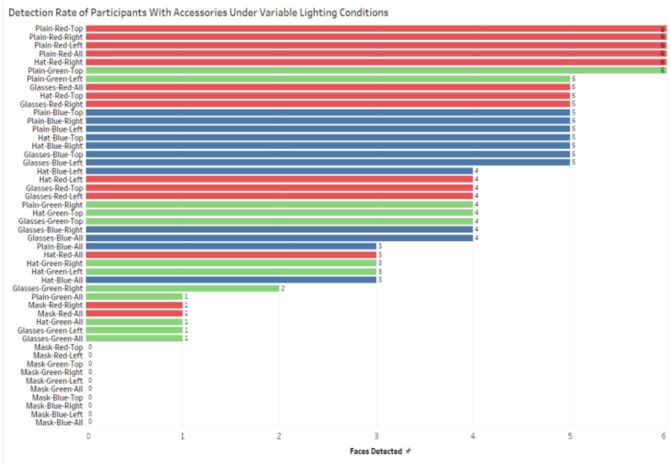

Fig. 3. Detection Rate of Participants with Accessories Under Variable Lighting Conditions

While there is a low number of participants within this testing, 48 different light scenarios/types of accessories were used per participant.

Through the observation of tests that were run, it has been seen that the camera used seems to mostly focus on mouth, nose, and eyebrows as its reference point while eyes not as much. The detection rate with glasses was relatively good but another test that could have been carried out would have been with an eye patch to see if eyes do really make a difference.

Fig. 4. Example of output showing recognition of a face, but not identity

3 Evaluation and Discussion

The initial tests suggest that a testing framework would be a useful tool, as we can see that differing conditions have an impact of the efficacy of the recognition system. Many inexpensive “AI” cameras indicate that they work under “good” light conditions without this being clear as to the exact meaning of “good”.

Under daylight scenarios, the camera performed as expected and facial features were accurately recognised, the colour green has been the most problematic for the camera while the colour red had the least impact on all results for different accessories worn. The tests were all carried out between 2:00pm and 4:00pm in the same room, However, it would be more accurate to control the variation of ambient light when testing the camera on participants. This could be achieved with a light meter. What has been learned from this project is that light really does affect the quality and the outcome of the data being collected. It has been noticed that taking pictures of participants at different times of the day would impact the results as pictures that were taken on cloudy or late hour days were more likely to cause the Huskylens to not detect a person during tests. It may be that the Huskylens uses an infra-red component of the light, but this is not specified in the documentation.

This work has been limited by a low number of subjects and could be improved with a wider range of ages and ethnicity. Different age scales for genders could have resulted in bias of the data results which could be the reason of the camera being more likely to detect males than females as males were within a similar age in the test whilst there was greater variation in the ages of the female participants. With a greater number of subjects, it would be possible for example to calculate an accurate recognition rate or other appropriate metric for “AI” Camera systems [15], and a potential user would use the metric best suited to their application.

The approach used could easily be adapted to ensure the subject was within the one metre range and a servo-motor could provide some small changes of camera angle.

Considering that some people use eyebrow pencils to draw or change the shape of their current eyebrows and use different coloured lipstick, this could become part of the test as the camera could potentially consider the participant as another person if their facial features or other parts surrounding the face were to slightly change.

It would also be appropriate to see if a camera could detect a photograph of a face, as it is important to ensure that a subject is present and not just their image.

There is a growing awareness of a need for transparency in AI systems, and this extends to accurate and detailed specifications that will allow users to determine if the equipment is fit for the intended task. Using a standardized test framework would assist with this, as it is independent of the Face Recognition technology used.

References

1. Shabbir, J. and Anwer, T. ‘Artificial Intelligence and its Role in Near Future’, *arXiv:1804.01396 [cs]* [Preprint]. Available at: http://arxiv.org/abs/1804.01396 (2018) (Accessed: 29 December 2021).

2. Donepudi, P.K. ‘Application of Artificial Intelligence in Automation Industry’, *Asian Journal of Applied Science and Engineering*, 7(1), p. 15. (2018).
3. Kaur, P. *et al.* ‘Facial-recognition algorithms: A literature review’, *Medicine, Science and the Law*, 60(2), pp. 131–139. doi:10.1177/0025802419893168. (2020).
4. Labati, R.D. *et al.* ‘Biometric Recognition in Automated Border Control: A Survey’, *ACM Computing Surveys*, 49(2), pp. 1–39. doi:10.1145/2933241. (2016).
5. Sharif, M., Javed, M.Y. and Mohsin, S. ‘Face Recognition Based on Facial Features’, p. 8. (2012).
6. Chen, Y.-L. *et al.* ‘Accurate and Robust 3D Facial Capture Using a Single RGBD Camera’, in *2013 IEEE International Conference on Computer Vision*. Sydney, Australia: IEEE, pp. 3615–3622. doi:10.1109/ICCV.2013.449 (2013)
7. Cong, R. and Winters, R. *How It Works: Xbox Kinect*. [online] Jameco.com. Available at: https://www.jameco.com/jameco/workshop/howitworks/xboxkinect.html (2019). [Accessed 12 Jan. 2022].
8. Sunil, A. *et al.* ‘Usual and Unusual Human Activity Recognition in Video using Deep Learning and Artificial Intelligence for Security Applications’, in *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*. Erode, India: IEEE, pp. 1–6. doi:10.1109/ICECCT52121.2021.9616791. (2021).
9. Adjabi, I. *et al.* ‘Past, Present, and Future of Face Recognition: A Review’, *Electronics*, 9(8), p. 1188. Available at: https://doi.org/10.3390/electronics9081188. (2020).
10. Kortli, Y. *et al.* ‘Face Recognition Systems: A Survey’, *Sensors*, 20(2), p. 342. Available at: https://doi.org/10.3390/s20020342. (2020)
11. Lal, M. *et al.* ‘Study of Face Recognition Techniques: A Survey’, *International Journal of Advanced Computer Science and Applications*, 9(6). Available at: https://doi.org/10.14569/IJACSA.2018.090606. (2018)
12. Burgess, P. *Adafruit NeoPixel Überguide*. [online] Adafruit Learning System. Available at: https://learn.adafruit.com/adafruit-neopixel-uberguide/arduino-library-use. (2013).
13. The Arduino Uno. https://docs.arduino.cc/hardware/uno-rev3 (Accessed: 20 June 2022).
14. *Gravity: HUSKYLENS - An Easy-to-use AI Machine Vision Sensor*. Available at: https://wiki.dfrobot.com/HUSKYLENS_V1.0_SKU_SEN0305_SEN0336. (Accessed: 6 June 2022).
15. Sundaram, M. and Mani, A. ‘Face Recognition: Demystification of Multifarious Aspect in Evaluation Metrics’, in S. Ramakrishnan (ed.) *Face Recognition - Semisupervised Classification, Subspace Projection and Evaluation Methods*. InTech. Available at: https://doi.org/10.5772/62825. (2016)