VICTORIA UNIVERSITY OF WELLINGTON TE HERENGA WAKA
Wellington Faculty of Engineering Te Wāhanga Ahunui Pūkaha
Ethics and Sustainability Symposium

Facial Recognition Technology Code of Ethics

Michaiah Thoms

Abstract

Facial recognition technology (FRT) is a tool used to identify individuals by scanning their face and comparing it to a database of facial images. In the last decade, the use of FRT has rapidly increased and is now used all around the world. This has resulted in many ethical issues caused by taking an individual's images without consent to train FRT programs, FRT misidentifying individuals, and using FRT to classify individuals as part of specific groups based on appearance alone. Because of these issues, a new code of ethics is proposed to provide a guideline for the creation and use of FRT. The code of ethics will prevent images of individuals from being taken and used without their consent and prevent FRT from being used unethically.

Keywords: FRT; Facial recognition; Code of ethics; Ethical framework.

1. Introduction

Facial recognition technology (FRT) is a method of automatically identifying individuals by having a program compare an image of their face against a database of other facial images [1,2]. FRT programs find and compare features of each face, such as eyes, nose and mouth, and use that information to find other faces similar to the one being scanned [1,2]. However, FRT programs need to be taught how to identify facial features, which means they need to train on often thousands or even millions of facial images before they can reliably find a face in other images [2,3,6]. If an image and details of an individual are available on a database, an FRT program should be able to identify that individual in other images of their face, even if the program has never seen those images before [1,2].

It is believed that FRT was invented in the 1960s by Woodrow Wilson Bledsoe, but it has only seen significant use in the last decade [1]. FRT has seen a large range of uses – identifying dead bodies, enhancing international airport security, preventing impersonation, finding missing people, tracking criminals, authentication in smart devices, reading emotions, and some medical studies have even used FRT to diagnose consenting patients [1-4]. Unfortunately, FRT has caused and been used in many unethical actions. Such as - taking photos of individuals without their consent to train FRT programs, or using FRT to identify, surveil and sometimes even persecute protestors and members of religious minorities [2,3,5,6].

1.1 Objective

This paper aims to analyse the ethical issues caused by FRT and what is being done about them. Then propose a code of ethics that would provide a guideline for those who develop and use FRT so they can avoid further ethical infringements. The proposed code of ethics is applied and evaluated against a specific case of FRT ethical infringement.

2. Literature review

The following literature review will analyse literature of the last three years, looking specifically at ethical issues caused by FRT and what is being done about them. A code of ethics for FRT will then be created using the information gained from this analysis.

2.1. Collection and distribution of facial images without consent

As FRT programs need to be trained on often thousands or even millions of facial images, those who develop FRT programs often gain those images unethically, without the consent of the individuals in the images [2,3].

Cameras in public locations are one of the primary sources of the enormous quantity of images needed for training [3]. In 2015, scientists at Stanford University in California took 12,000 images of individuals from a webcam in a café that was being live-streamed online [3]. In 2016, researchers at Duke University in Durham, North Carolina took 85 minutes worth of footage of students walking around their campus, equivalent to 2 million video frames [3]. In both cases, the images taken were published online for training FRT programs, and none of the individuals in the images gave their consent or were made aware of this happening [3].

Another method of acquiring these images that is more common is scraping them from the internet, generally social media sites [5,6]. As people all around the world post millions of images of themselves and others online every day, it is a simple matter to collect as many of those images as required and use them to train FRT programs. Many universities and companies have done this in the past and put together databases with billions of images that could be used for training FRT programs and made them publicly available [5,6]. And yet again, none of the image's owners or the individuals in the images gave their consent to this mass collection of their images for public use in FRT programs [5,6].

2.2. Use of inaccurate facial recognition technology

There are also major ethical issues with how FRT is used, particularly in law enforcement where it is used to identify and find criminals [2,5]. The main issue is that FRT commonly misidentifies individuals as criminals on a watchlist, and heavily relies on humans to confirm its identifications [2]. This is primarily caused by FRT programs being used without having their accuracy properly measured or while they are known to be extremely inaccurate and/or biased [2,5].

In January 2020, Robert Williams was arrested after the Detroit police's FRT program misidentified him as a watch thief, after comparing his driver's license photo to the blurry surveillance footage from the store that was robbed [2]. Williams was detained for 30 hours. During that time, Williams claims the image of the thief looked nothing like him, but the detective stood his ground because the FRT program said it was him [2]. Later, after Williams was released, and a complaint was filed, the Detroit police chief James Craig acknowledged the FRT program, by itself, was wrong 96% of the time [2]. FRT has also proven to be much less accurate at identifying females and individuals with darker skin, such as Robert Williams and the thief were in this case [2,5]. It is highly unethical to identify individuals as criminals like this when the FRT system used to do so is correct only 4% of the time.

There are also records of police in the US and UK using FRT to scan crowds looking for criminals on their watchlists [5]. This is highly unethical when you consider that FRT was proven to be much less accurate at one-to-many identification like this compared to one-to-one identification like the Williams case [2]. There was also a case in the UK where a man sued the police in South Wales, claiming his privacy had been breached when he was scanned in a crowd with FRT [5]. The court ultimately ruled against him, but in the process uncovered that the police hadn't sufficiently checked the accuracy of their software [5].

2.3. Using FRT to identify, classify and surveil individuals

There are also major ethical concerns with how governments could use FRT, given its ability to identify and classify individuals based on their age, gender, sexual orientation, race, nationality, religious beliefs, or disabilities [3,5]. Governments could use FRT to identify and suppress/persecute their opposition or protestors or limit the freedom of the general population in their countries in many other ways [5].

This has already been seen with China's use of FRT to identify and persecute Uyghur people (a predominantly Muslim minority ethnic group) from other ethnicities [2,3,5,6]. The Chinese government was using FRT in surveillance cameras to find and detain Uyghurs on mass and put them in detention camps, which the government claimed were re-education centres aimed at quelling a terrorist movement [2,3,5,6].

The governments of Russia, China, India, and South Korea are also using FRT to trace COVID-19 contacts and enforce quarantine [5]. In March 2022 Vladimir Bykovsky, a Moscow resident, had recently returned to the country and was undergoing a 2-week quarantine [5]. At one point during his quarantine, he left his apartment for a moment to throw out his rubbish, where a camera using FRT saw him and 30mins later police arrived to give him a fine and a court date as he violated his quarantine [5]. Researchers are worried that this use of FRT may not end with the pandemic and could result in a loss of freedoms for society [5].

2.4. Current attempts to address the problem

Given the major ethical issues around FRT, there have been numerous attempts to restrict the use of FRT and rectify existing issues with the technology. Many surveys have been taken from the public and artificial intelligence researchers regarding FRT. The majority of those surveyed have major concerns with FRT and generally don't trust it to be used ethically [3-6]. The majority also agree that it should only be used when there is a clear public benefit and with the informed consent of those it affects [3-6].

In September 2019, four researchers respectfully asked that a study on training FRT algorithms to identify Uyghurs be retracted [3]. In health research, all identifying information is removed from facial imaging used in FRT to protect the privacy of the individuals in the images [4]. Scientists have made many suggestions and requests for laws that prevent the collection of facial images from public places, temporary bans on the use of FRT until better restrictions are in place, and a US federal office that manages FRT applications [5]. Unfortunately, scientists alone cannot enforce these suggestions, but they can and are campaigning loudly for them and similar ideas [6].

3. Code of ethics/sustainability

The following are the proposed principles for a code of ethics that would regulate the creation and use of facial recognition technology. These principles have been put together based on the ethical issues and attempts to fix them identified in the above literature review.

3.1. Attain informed consent for collection and use of facial images

Avoid collecting or using images of an individual's face without their knowledge and informed consent. The informed consent must clearly communicate any intended use of the individual's facial image. Informed consent must be re-acquired before an individual's facial image is used for any purpose not originally communicated to the individual.

3.2. Allow individuals to update or delete their facial images

An individual should be able to update or delete any facial images collected from them at any time.

3.3. Securely store collected images and don't distribute them without consent

Collected images should be stored securely so that they can only be accessed by authorised individuals. Collected images should not be accessible by or distributed to other parties unless consent is acquired from the individuals that are in the images.

3.4. Maintain a high accuracy for facial recognition technology regardless of an individual's appearance

FRT should be able to identify individuals consistently and accurately and an individual's age, gender, sexual orientation, race, nationality, or disabilities should have little to no impact on the accuracy.

3.5. Do not use facial recognition technology to classify individuals

FRT should not be used to classify individuals based on their age, gender, sexual orientation, race, nationality, religious beliefs, or disabilities.

3.6. Provide alternative means of identification

Whenever FRT is used to identify an individual, a reasonable secondary method of identifying the individual must be taken to confirm the identification.

4. Case study discussion

In the last few years, the New Zealand Police have experimented with FRT and incorporated it into its existing systems [7-10]. During this process, they have had a few ethical issues with the use of FRT [7,8]. Therefore, the proposed code of ethics will now be applied and evaluated against the NZ Police's use of FRT.

4.1. Attain informed consent for collection and use of facial images

Attaining informed consent for the collection and use of facial images is difficult when such large numbers of images are required for FRT to be effective. The NZ Police have tried a few unethical options already, such as using Clearview AI which is a database of around 3 billion images scraped from social media and other public websites [7]. Officers in NZ have also gone around taking photos and other details of individuals in public places [8] and have also used surveillance cameras to collect facial images [9]. The NZ Police are committed to maintaining transparency around their use of FRT [10], so individuals would at least be aware of how their images are being used.

To follow this principle, the NZ Police would have to get informed consent from everyone they collect images from, which would be simple when taking images of those in public places. Alternatives could include an opt-in system when individuals get photographed for their driver's license, where they get asked if their image could be used to enhance the police's FRT systems.

This principle is severely limiting and requires a significant amount of effort and resources from those using FRT. Therefore, this principle is, unfortunately, the hardest to follow and, realistically, will be ignored in most cases. However, it addresses one of the most crucial ethical aspects of FRT.

4.2. Allow individuals to update or delete their facial images

Once NZ Police collect facial images of individuals with their informed consent, the individuals should then be provided with a method of updating or removing their image on the NZ Police's database. An extensive amount of time and resources is required to set up an entire system that would allow individuals to do this. However, this would be ultimately beneficial for both parties, as individuals have control over their images, and if individuals keep their images updated, the NZ Police wouldn't need to worry about updating the images manually.

Ultimately, this principle is probably the least important, as once an individual has given their informed consent for their image to be used, they are unlikely to want or need to change it or remove it from a database. It is also unlikely for any major ethical issues to arise from not following this principle.

4.3. Securely store collected images and don't distribute them without consent

The NZ Police claim they keep collected images stored securely per NZ's Privacy act [9]. Nothing is mentioned about not distributing the images to other parties, but it is unlikely that the NZ Police would ever need to. However, if they ever decide to distribute collected facial images, they should first attain the informed consent of the individuals in the images. Given that this is unlikely to occur, this principle would have very little impact on the NZ Police, while ensuring the security of the personal information of the individuals in the database.

4.4. Maintain a high accuracy for facial recognition technology regardless of an individual's appearance

Given FRT's potential use by the NZ Police to find and identify criminals, this is an important principle to consider, especially as the population of NZ has a diverse range of ethnicities. The NZ Police are aware that the accuracy of their FRT systems is important and are currently not using FRT on live cameras due to the negative impact on the accuracy and bias of FRT on live feeds [7]. The NZ Police should be doing regular checks to ensure that their FRT systems are accurate and unbiased so that they can avoid misidentifying individuals as criminals or vice-versa and can accurately identify individuals of any appearance.

This principle is perhaps the most important in this code of ethics, especially for law enforcement as it could erode the public's trust in their effectiveness. This is due to the potential for major consequences of misidentifying individuals e.g., arresting the wrong person or letting criminals run loose when they could've been identified by a more accurate FRT system.

4.5. Do not use facial recognition technology to classify individuals

For NZ Police this principle would likely not come up often as their FRT systems would be generally focused on finding and identifying specific individuals. However, there is potential for FRT to be used to identify individuals as members of gangs or terrorist groups based solely on their appearance. If or when these situations arise, the use of FRT in this manner should be avoided and only used to find known members of specific groups.

This principle is not overly limiting but does prevent a dangerous use of FRT that could cause major ethical issues, making it a key principle in this code of ethics.

4.6. Provide alternative means of identification

Whenever the NZ Police use FRT to identify individuals, they should always follow it up with a secondary means of identification. This could involve getting a human to compare the images, as done by police in the UK [2], or by asking to see the individual's driver's license (or other forms of identification) and comparing it to the record of the individual they are looking for.

This principle ensures that FRT isn't solely relied on as a means of identification as FRT can produce incorrect results periodically. Confirming identities through secondary verification measures provides a safeguard against ethical issues caused by FRT misidentifying individuals.

5. Conclusion and recommendations

Facial recognition technology has the potential to be an extremely useful tool with a large range of uses, but its creation and use have caused many major ethical issues. For small scale applications, such as using FRT in one-to-one authentication in smart devices, there are little to no issues with its use. But in large scale uses such as by governments and law enforcement, FRT is very difficult to use without causing ethical issues.

The proposed code of ethics in this paper provides a guideline for the ethical use of FRT. It ensures that individuals have control over how their facial images are being used and it reduces the potential of FRT being used unethically. It is highly recommended that all who develop and use FRT follow this proposed code of ethics.

References

[1] D. Dharaiya, "History of Facial Recognition Technology and its Bright Future," ReadWrite, 12 March 2020. [Online]. [Accessed 26 April 2022].

[2] D. Castelvecchi, "Is facial recognition too biased to be let loose?," *Nature (London),* vol. 587 (7834), pp. 347-349, 2020. Available:https://doi.org/10.1038/d41586-020-03186-4

[3] R. Van Noorden, "THE ETHICAL QUESTIONS THAT HAUNT FACIAL-RECOGNITION RESEARCH," *Nature (London),* vol. 587, no. 7834, pp. 354-358, 2020. Available: https://doi.org/10.1038/d41586-020-03187-3

[4] "A survey of US public perspectives on facial recognition technology and facial imaging data practices in health and research contexts," *PloS one,* vol. 16, no. 10, pp. e0257923-e0257923, 2021. Available: https://doi.org/10.1371/journal.pone.0257923

[5] A. Roussi, "RESISTING THE RISE OF FACIAL RECOGNITION," *Nature (London),* vol. 587, no. 7834, pp. 350-353, 2020. Available: https://doi.org/10.1038/d41586-020-03188-2

[6] "Facial-recognition research needs an ethical reckoning," *Nature (London),* vol. 587, no. 7834, pp. 330-330, 2020. Available: https://doi.org/10.1038/d41586-020-03256-7

[7] P. Pennington, "Review prompts police to halt plans to use facial recognition technology," RNZ, 9 December 2021. [Online]. Available: https://www.rnz.co.nz/news/national/457588/review-prompts-police-to-halt-plans-to-use-facial-recognition-technology. [Accessed 2 May 2022].

[8] T. A. Hurihanganui, "Police using app to photograph innocent youth: 'It's so wrong'," RNZ, 26 March 2021. [Online]. Available: https://www.rnz.co.nz/news/in-depth/437944/police-using-app-to-photograph-innocent-youth-it-s-so-wrong. [Accessed 2 May 2022].

[9] P. Pennington, "Police setting up $9m facial recognition system which can identify people from CCTV feed," RNZ, 31 August 2020. [Online]. Available: https://www.rnz.co.nz/news/national/424845/police-setting-up-9m-facial-recognition-system-which-can-identify-people-from-cctv-feed. [Accessed 2 May 2022].

[10] C. Jones, "An informed use of facial recognition technology by NZ Police," 11 June 2021. [Online]. Available: https://data.govt.nz/blog/police-frt/. [Accessed 2 May 2022].