Marisa Shuman and Marieke Thomas

Bias in Facial Recognition

CSCI 77800 - Ethics and Computer Science

Fall 2022

Final Project

Table of Contents:

Introduction:

Michael Oliver, Nijeer Parks and Robert Williams are three men living in the United States that do not have much in common. Robert Williams is from Detroit, Michigan, Nijeer Parks from New Jersey, and Michael Oliver is from Ferndale, Michigan; all are different ages and with different occupations. They do however have two things in common: all three are Black men and all three were misidentified using facial recognition and arrested for a crime they did not commit (Johnson 2022). While ultimately cleared of their separate and unrelated charges, all three had to mount a defense against cases that were built on the fact that a computer matched their face with that of a criminal. In this paper, we will examine the technology that led to these arrests and how it works. We will explain the significant biases that currently exist in facial recognition technology and identify the many benefits and significant harms that it poses to our society. Finally, we will present policy recommendations addressing the question: given that we know this technology has flaws that can be solved, should it be used in its current state in our world?

Background Information:

Facial recognition is a type of artificial intelligence that is currently being used in a multitude of ways and situations around our world. This constantly developing technology is used on a daily basis by the majority of people in the United States simply when they try to unlock their phone. Biometric phone unlocking however is just the tip of the iceberg. Facial recognition is being used in airports and security, law enforcement, tracking individuals in protests, other surveillance, and was formerly used in tagging Facebook images. While the importance of these uses spans a large range, it is essential to note that we are moving towards an increasingly technological world and that facial recognition is not going to disappear.

Currently, China is paving the way with their facial recognition implementation. The Chinese government has installed an immense network of cameras throughout the country that track individuals as young as nine years old (Ng, 2022). This tracking is used to prevent crimes, identify criminals, and help try to modify the behavior of all citizens. While this technology is not being used in the United States on such a wide scale, there is the potential to do so. Currently, there is an ongoing debate within the U.S. about whether the government should have the ability to use this technology to keep us safe.

Before determining the value of this technology, it is essential to understand how it is trained and how it works. According to Gupta and Mallick, Facial recognition is trained through machine learning, a process which essentially trains a computer based on pattern recognition (2022). Training data is classified and fed into the computer, which then creates a model of the categories. This is often done using Convolutional Neural Networks (CNNs), which are designed to pick out which features that distinguish between images. In our very simple facial recognition model, for example, we trained the computer with data sets of male and female pictures. When the model is being trained, the computer takes each of the input images and converts it into an array of numbers. The computer then embeds each of the images into a feature space, essentially plotting each image as a point in a graph with many dimensions. At first, the images are likely to appear somewhat randomly in this feature space. The computer iterates through the data, repeatedly modifying the feature space used so that images of the same category are brought closer together and images of a different category are separated further.

In our Teachable Machines example, the computer goes through 50 iterations, and on each iteration it attempts to make two separate clusters of images, one of women and one of men. One of the benefits of modern machine learning as opposed to earlier versions of artificial intelligence is that the computer itself decides mathematically which image features are important, rather than the programmer needing to hard-code the key features. According to AWS, earlier versions of facial recognition AI (many still in use today) use hard-coded metrics such as the distance between eyes, depth of eye sockets, contour of the lips, etc rather than a neural-network-generated feature space.

Per Gupta and Mallick, once the computer is trained based on the training set, it can now classify new images. When the computer is exposed to a novel picture, it classifies the new picture using the model that it developed, and then compares that picture to the preexisting categories. There are several different algorithms which can be used to make this comparison, ranging from accurate but inefficient (such as comparing the new image to every existing image) to more efficient solutions (such as the k-nearest neighbors algorithm, which finds a certain number of the nearest points and classifies the new image as whatever the most prevalent category among these neighbors is). In our simple facial recognition example, we are classifying each new image into one of only two categories, so a good algorithm to use would be Support Vector Machines, which creates a vector essentially representing the average of each category and then compares the new image to each of these category vectors. In a more complex facial recognition system, the goal is not merely to detect if a face is male or female, but to determine the actual identity of the face. In that type of system, each person would be a different category, so there could be billions of different categories, and the k-nearest-neighbors algorithm would be preferable.

When training machine learning algorithms, it is essential to use an enormous quantity of data; otherwise, random differences in the data sets can lead to erroneous classifications based on extraneous information. For example, imagine a model that is trained on just two pictures, one of Marieke and one of Marisa. Rather than detecting differences in our facial features, the computer would likely create a model that focuses on more obvious differences such as shirt color, distance to the camera, and lighting levels. If we were to switch clothes, the model would likely misclassify us. Similarly but on a larger scale, facial recognition algorithms are often created from data sets that disproportionately contain images of white and Asian faces, resulting in an algorithm that is less accurate at classifying black and Latino faces. For example, one study found that commercial gender classification systems were far more accurate at classifying light-skinned males (with an error rate of 0.8%) than dark-skinned females (with an error rate up to 35%) (Buolamwini & Gebru, 2018). A different study found that facial recognition software created by Western companies classified white faces more accurately than Asian faces, while software created by East Asian companies classified Asian faces more accurately than white faces, most likely due to biases in the underlying data used to train the algorithms (Phillips et al., 2011). In order to address these persistent biases, the training data used to create facial recognition algorithms must be expanded and diversified, specifically including more black and latino faces.

In the next section, we will explain how we modeled bias in this technology.

Facial Recognition Example Project

In order to illustrate the pitfalls of training facial recognition software with a biased data set, we created two primitive facial recognition models using Google's Teachable Machines API. This program contains a pretrained neural network and allows users to essentially train the last layer of the neural network. Users set their own categories and upload images into each of those categories. In this case, we set categories of "Male" and "Female" and uploaded photographs from the UTKFace Open Source face data set, which we sorted using glob patterns, which is a more limited version of regular expressions, as described in Appendix 2. For our biased algorithm, we uploaded 2,329 images of only white men and women, while for the less biased algorithm, we added images of multiple ethnicities for a total of 5,229 images. The Teachable Machines program trains the neural network, feeding each image through the model 50 times in order to train the AI to accurately distinguish between the two data sets.

After training the AI twice to create biased and less biased facial recognition programs, we uploaded these two programs to our website so that users can test their own face with both programs. Of course, one user will not be able to test how the algorithms respond differently to different ethnicities, so we also tested a series of photos of celebrities against the two different algorithms (uploading the photos directly to the programs that we created on the Teachable Machines website, rather than via the API). Both programs correctly identified the genders of the vast majority of celebrities of all races; however, the biased program failed to correctly identify the genders of Serena Williams, Jimin, and Bruno Mars, while the less biased program did correctly identify all three. On our website, we included the results for these 3 celebrities, along with Kate Hudson, acting as a control because she is a white woman who has short hair and is sometimes considered to look masculine. Both programs correctly gendered her.

Using Google's Teachable Machine API was a great way to demonstrate the bias that is undeniably linked within the machine learning process of facial recognition. Unfortunately, we also recognize that there were a few limitations that we could not control. First, we used over twice the amount of images when training our less biased algorithm. As a result, this procedure conflates the effects of having a more diverse sample with simply having more sample images. A more scientific approach would be to use the same number of images in both cases. Additionally, Google's Teachable Machine requires specific categorization (as does all Facial Recognition) and as a result we used the categories of male and female. It is imperative to note that gender is not actually binary and in reality is a spectrum.

Our project is an example of how biased training sets produce biased results in facial recognition software. Of course, a commercial facial recognition program would be trained on far more than a few thousand images, and rigorous testing of the resulting biases would be automated with thousands of images as well. Because our model is trained on such a small and deliberately-biased number of images, the biases are exaggerated. Nonetheless, the same biases are documented to exist in commercial facial recognition systems (Buolamwini & Gebru, 2018). In particular, as with Serena Williams in our example, black women are often disproportionately categorized as male, while white people are correctly-categorized much more frequently than people of other ethnicities. When biased data sets are used to train neural networks, the resulting AI is biased.

For a link to our website and an explanation of the JavaScript file, see Appendix 1.

Given that facial recognition software has known and documented biases, there is a fierce debate over whether this technology is ready for use. In the next two sections, we will examine the arguments in favor and against the usage of this technology.

Arguments in Favor:

In this section, we will discuss the benefits of using facial recognition software. The primary benefit of using facial recognition software is that it is convenient and useful. For example, many people use facial recognition to securely unlock their phone or access important accounts such as their banking app (Apple 2022). While facial recognition alone is still less secure than a passcode, when it is used in conjunction as a form of 2-factor authentication, it can help keep sensitive accounts more secure ("5 Most Common Uses of Facial Recognition," 2022). The chief argument for allowing the use of facial recognition software in these types of circumstances is that the benefits outweigh the drawbacks. Users who opt in to using facial recognition to access phones and accounts are voluntarily choosing to do so, and those with privacy concerns can choose not to. As this is an opt-in system, each individual can decide for themself if they feel that the convenience of facial recognition makes it worth using. The simple fact that the public can often elect to use facial recognition is a strong argument for its use as banning facial recognition in these circumstances would be paternalistic government overreach.

In another common application of facial recognition software, Facebook has employed the technology to detect the faces of people who appear in photographs and automatically suggest tags, saving users time and making it easier to share photos with friends and family (Pesenti, 2021). While allowing users to save a few seconds may not be very compelling, Facebook has also used facial recognition to make the site more accessible to all users by incorporating it into Automatic Alt Text, which is a program that automatically generates descriptions of images for people who are visually impaired. Incorporating facial recognition software into the program allows users to know who is in the photographs that appear in their feed, allowing them to interact more fully and independently with the site, which is a key goal of accessibility features. While there are some privacy concerns around allowing social media to tag users without their explicit permission, it can be argued that once their face appears in a picture, anyone looking can already identify them. Tagging pictures does not change who is in the picture, it simply allows visually impared users to know who is there, just like everyone else already can. In short, facial recognition can be used to make sites easier to use for everyone, especially those with special needs. In addition, without automatic tagging, it is common for someone to post a picture online of another person without their knowledge or consent. Facial recognition software can alert users when other people post photographs or videos of them, allowing people to preserve their privacy by controlling where and when they appear online. In the future, social media sites could implement a feature where users can take down a picture featuring them, or even a feature where everyone in a picture must consent in order for it to appear publicly. Without automatic tagging using facial recognition, users may not know when someone else posts a harmful picture of them – but even if they are not aware, the picture is still there. Automatic tagging does not change who is in a picture, but it allows those people the knowledge and power to control how their image appears online.

In addition to making technology easier and more convenient, facial recognition has been used by law enforcement to investigate and prosecute crime. For example, according to the New York Police Department's website on Facial Recognition, they use the software to generate leads on cases by comparing photographs of perpetrators (such as from security camera footage of a robbery) to a database of arrest photographs. In 2019, for example, this software identified 2510 possible matches. This software can rapidly identify potential suspects, allowing for investigations to proceed quickly, before key evidence is displaced, and freeing officers from tedious searches through photo databases, leaving them more time to conduct thorough witness interviews, investigate potential leads, and gather evidence. Of course, facial recognition software has its limits– the NYPD explicitly states that "a facial recognition match does not establish probable cause to arrest or obtain a search warrant" but rather is used to generate leads that then must be investigated further. In this context, there is a huge amount of human oversight, as is necessary and appropriate. While facial recognition has led to wrongful arrests, those wrongful arrests are ultimately the responsibility of the officers making them. Officers who fail to follow proper procedure, for example by failing to verify if the suspect has an alibi that makes it impossible for them to have committed the crime, should be reprimanded. However, disallowing the use of a technology that is a helpful tool in solving crimes simply because some officers misuse it is counterproductive; the solution to officers misbehaving is to punish those officers, not the technology.

The unfortunate reality is that with or without facial recognition AI, wrongful arrests and convictions are depressingly common in this country. The Georgia Innocence Project estimates that "4-6% of people incarcerated in US prisons are actually innocent" (2022). Eyewitness testimony has been repeatedly shown to be biased and unreliable. For example, among people who have been wrongfully convicted and later exonerated due to DNA evidence, 70% of them were originally convicted based on witness testimony (Bohannon, 2015). Pseudoscientific evidence, such as analysis of bite marks, blood stains, or shoe prints, is frequently used in court despite "substantial rates of erroneous results" (2008). For example, a study in which dentists were asked to classify marks found that in the vast majority of cases, they could not even agree on whether or not the mark was a bite mark, let alone whether it matched the suspect's teeth (Freeman & Delger, 2020). Given this context, the bar for determining whether or not to allow police to use facial recognition software should not be "does this technology ever lead to wrongful arrests?" Instead, we must ask, "Does this technology lead to *fewer* wrongful arrests than eyewitness testimony and the other biased methods currently used by law enforcement?" While facial recognition accuracy varies based on image quality, it has been found to be roughly 90% accurate when used on pictures of people taken in public (Chowdhury, 2022), which is of course not ideal but still wildly more accurate than witness testimony and forensic “evidence” used in courts today. Facial recognition has well-known biases which we have described above, but it is likely still less biased than eyewitness testimony, and when used with additional safeguards and common-sense policing, can give additional information to police. Police must be trained on how to use this tool and must understand the level of error that it involves; nonetheless, it provides valuable information that should not be ignored. Every time that facial recognition correctly identifies the perpetrator of a crime makes it less likely that police will accuse an innocent person.

One of the most important uses of facial recognition by law enforcement is locating victims of child trafficking (Simonite, 2019). The Thorn Spotlight program scans through pictures of online sex ads, checking for matches against photographs of missing children. In the last four years, this program has assisted in identifying 17,092 children and teens who had been victims of sex trafficking (Thorn, 2022). Without this technology, these results simply would not be possible. An AI algorithm can scan through millions of images in seconds, when it would take human investigators weeks, or they might never notice a hit. This is essential time that victims cannot afford to wait. Once victims are safe, the program can also be used to help prosecutors build a case against the perpetrators, for example by quickly locating ads placed 6 months previously (Hartsock & Woods, 2018). Of course, this AI does not replace human investigators. In these types of cases, the key involves building trust with victims over time, and a facial recognition algorithm cannot do this. It can, however, conduct online searches rapidly and efficiently, freeing investigators from tedious grunt work and allowing them to focus their efforts on building relationships and convincing victims to accept help.

One common criticism of facial recognition by law enforcement in public places is that it is an invasion of privacy. However, the counterargument is that when someone is in public, they have already put themselves in a space where others can see and recognize them– using facial recognition is not substantially different from if they happened to run into someone they know. Anyone who is in public should already be operating under the assumption that their actions are visible and they could be recognized by those around them. If someone commits a crime that is caught on video, it is absolutely appropriate for them to be identified based on that video, just like someone who commits a crime in front of an eyewitness can be identified by the witness. Facial recognition is essentially just an extension of the methods law enforcement officers already use to solve crimes.

The final argument in favor of using facial recognition technology is that this allows us to detect flaws and improve them. This technology is already heavily used in other countries and by private companies– there is no way to put the cat back into the bag. Outlawing this technology would not actually prevent its use, but merely drive it underground, where it cannot be regulated or studied. Instead, by allowing its usage in controlled environments, we can uncover flaws, such as the bias against black people and women. Companies can then work to fix those flaws, such as by using more diverse training data sets.

There are a range of strong arguments for why facial recognition software should be used– it is useful and convenient, it can help those with visual impairment, it can be used to locate suspects of crimes and victims of human trafficking, and using it allows us to improve it and weed out biases. In the next section, we will consider arguments against its usage.

Arguments Against:

While it is undeniable that facial recognition technology is a powerful tool that has the potential to better our society, it also has the ability to harm. We have already listed many flaws in the current status of facial recognition technology and machine learning, including the misidentification of individuals, especially including people of color and women. While ideally the number of individuals misidentified will go down over time and with more refinement of machine learning algorithms, we must still consider the current impact on countless individuals across the United States and the world. It can and should be argued that if there is any flaw in this technology, especially when that flaw targets a specific group of people, the technology should not be used. The greater good and benefit to many should not be valued more than the disruption and discrimination against a single individual, especially when that individual is already facing systematic injustices.

Looking past the significant biased flaws in facial recognition algorithms, we must also consider how much easier facial recognition makes surveillance. Before the rapid advancement of cameras and facial recognition algorithms, there was a limit to the amount of surveillance that the government or law enforcement agencies could achieve due to the limitations of physical manpower. Facial recognition has the potential to create an environment where the government, or private companies, can track where people are at all times when in public, which would be a massive invasion of individual privacy. Many people argue that an increase in surveillance means an increase in safety due to a decrease in crime. Unfortunately, the results are mixed. A 2015 study done by Michael Prix found that additional surveillance cameras added to the Stockholm subway system helped reduce crime by approximately 25%. Conversely, the Metropolitan Transportation Authority installed 784 cameras in the New York City Subway in 2020 but noted an increase in serious crime, though it should be noted that this data is from the Covid-19 pandemic (Nessen, 2021). The question then must be asked: is the loss of privacy in our day to day lives worth a potential but not definite increase in security? Ultimately, until there is a proven correlation between increased surveillance and crime reduction, it should not be used.

Law Enforcement Agencies are also using facial recognition technology to help identify suspects in a crime or identify individuals who are not legally allowed to be in the United States. There have been many incidents where agencies have used facial recognition technology in order to identify individuals participating in protests. During the Black Lives Matter protests in the United States in 2020, multiple reports showed facial recognition being used on the participating public. According to *How Facial Recognition Technology Can Be Used at a Protest*, this data could be used to compile a list of individuals for future matching and surveillance. In Hong Kong, the public has taken steps to counteract the state surveillance in order to peacefully protest the Hong Kong and Chinese governments which has included wearing masks, lasers and other cloaking devices (Cuthbertson, 2019). Facial Recognition technology has been employed in peaceful protests before any sign of unlawful activity due to racial biases held by those in power. Additionally, the United States Immigration and Customs Enforcement has started using facial recognition and DMV images that allow them to track immigrants in the U.S. and is able to use the identities to track their addresses with utility bills (Alms, 2022). Unfortunately these examples prove that facial recognition is being used around the world and in the United States to continue systematic racism and inequitable practices.

On top of biased data being used to train facial recognition programs, use of these programs by law enforcement presents an additional bias. According to the New York City Police Department's information on their use of facial recognition, when they use facial recognition software to identify a suspect, they typically do so by matching a photograph of someone committing a crime (such as a still taken from security camera footage) to a police database of mug shots. Because this database does not contain images from the general public but only from people who have prior arrests, use of this software is likely to fail to locate first-time offenders, and raise false positives among people with a prior criminal record. This means that the same group of people, disproportionately low-income black men, are likely to be continually harassed by law enforcement over the course of their lifetimes. Since our current criminal justice system is heavily racially biased, the usage of mug shots as a comparison will cause any false positives to be racially biased as well, even on top of the biases embedded in the AI algorithms.

Beyond the heavy bias that is embedded in this technology and its use in law enforcement agencies, we must consider the potential privacy violations committed in order to create training data sets. In order to train facial recognition algorithms, millions of images need to be gathered and sorted. It is an immense task to gather such a large compilation of images and many sets have been curated by some of the big technology companies, including: Alphabet, Amazon and Microsoft. We are currently living in a time where our phones are our cameras and they rarely leave our side, enabling us to take more pictures than ever before. However, taking a picture for personal use and memories is very different from taking one to be used to train a facial recognition algorithm.

There has long been concern over where the images used in training sets are taken from and if the image owner has given consent (Keyes 2019). Since 2015, there have been clear instances of images being collected explicitly without a person’s consent or knowledge. From a 2015 data set published by Stanford University that took its pictures from a public webcam in San Francisco, to two million images taken of students walking around the Duke University Campus in 2016, often, we are unaware our photo is being taken and even more unaware it is going to be subsequently used to train a facial recognition algorithm (Noorden, 2020). Further, it has been recently discovered that the Facial Recognition Verification Testing program that is run by the National Institute of Standards and Technology also uses images from vulnerable populations, including: children who have been victims of child pornography, former convicts -- both alive and dead, and potential U.S. visa applicants (Keyes 2019). This is an extreme example of the United States Government taking advantage of vulnerable populations to further their agenda. People should have the right to control their own image and likeness, and companies building facial recognition software should be required to seek permission from and fairly compensate those whose images are used in training data. Without these safeguards, facial recognition technology as it currently exists is exploitative and should not be used.

It is evident that there are many arguments that could be made as to why facial recognition both should and should not be used in our society. Ultimately, the intrusion of the privacy of individuals, specifically those who are most vulnerable, and the bias that is threaded throughout are issues that are too large to be ignored. Countless individuals have their privacy robbed and are still misidentified using facial recognition technology.

We have examined arguments for and against the use of facial recognition software. In our final section, we will evaluate these arguments and describe our policy position.

Our Position:

Given the biases that are embedded into facial recognition algorithms, we must ask ourselves: should this technology be used today? We hold the position that this technology should be used for low-stakes private enterprise, but should not be used in its current state by law enforcement or government agencies. Furthermore, for companies that are permitted to facial recognition, the algorithms used should be probed for biases, and the images used in training data should be obtained ethically. We consider each of these points in turn.

Facial Recognition by Private Companies:

We believe that use of facial recognition by private companies for opt-in services should be allowed. Permitted uses would include allowing smartphone users to unlock their phones, bank clients to log in to their accounts, and social media sites to tag users using facial recognition. In all of these cases, users should have the option to opt in to using facial recognition, or opt out and still have full access to the service provided. Additionally, there should be complete transparency in any data that the companies are collecting in order to continue the training of their algorithms, as described below. Any individual who is faced with the decision to opt in or out of the use of facial recognition should be fully informed of any potential use of their biometrics.We believe that the benefits to users of increased convenience and accessibility outweigh the risks of using this technology. We also feel that individuals have the right to decide for themselves if they wish to use facial recognition software on their own likeness. As long as users are knowledgeable of the use of facial recognition and retain the right to opt out of having it applied to them, we feel that the privacy concerns are minimal and this usage should be permitted.

Facial Recognition by Government and Law Enforcement:

We feel that facial recognition in its current state should not be used by government or law enforcement. While this software has the potential to help solve crimes and locate victims, we nonetheless feel that it gives too much power to law enforcement officers and has too great a potential for abuse. We know that at present, this AI is biased and inaccurate; its use by law enforcement will widen racial disparities in justice and lead innocent people, most likely black people, to be wrongfully arrested. In an ideal world, this technology could be used by officers who would fully understand its limitations and only arrest suspects if there was substantial additional evidence that they had committed a crime, and officers who made unwarranted arrests would be reprimanded. However, in the real world, this technology can be and already has been used to arrest people without sufficient evidence and even in spite of solid alibis proving that they could not have committed the crime. Police have a long history of arresting their first suspect even when faced with strong counter-evidence; this technology would make it even easier for police to do so. In other contexts, such as scanning public camera footage for wanted suspects or undocumented immigrants, use of facial recognition software contributes to a surveillance state that infringes heavily on individual privacy. We believe that the substantial benefits of using facial recognition software by government agencies do not outweigh the overwhelming risks to privacy and freedom from false imprisonment, and this software should not be used for that purpose.

Recommended Regulations

While we believe that use of facial recognition algorithms should be permitted in some circumstances, we feel that its use should be regulated. Some states have already taken steps to regulate its use. For example the Biometric Information Privacy Act, or BIPA that was passed in the Illinois legislature in 2008 (*Biometric Information Privacy Act (BIPA)*, 2022). BIPA protects the biometric data of an individual and prohibits its collection and use unless a company alerts the individual in writing about their intents, purpose and storage and receives written consent. Sadly, only three states currently have legislation in place to protect an individuals' data and privacy. With the lack of governmental constraints and regulation, other regulation must be put in place. The Future for Privacy Forum, a non-profit organization that focuses on privacy and data practices, created a list of "Privacy Principles for Facial Recognition Technology in Commercial Applications" in 2018. This suggested set of guidelines states that in order to create safe and effective facial recognition technology that protects the privacy and data of the population, it must include consent, respect for context, transparency, data security, privacy by design, integrity and access, and accountability.

We believe that policies such as BIPA or the Privacy Principles for Facial Recognition should be adopted on a national level. Having a set of standardized federal or global regulations would ensure that all facial recognition technology is moving forward with the population and privacy in mind. Additionally, the algorithms should be monitored for racial and other biases, and companies making them should be required to take steps to counter any biases detected. Individuals should have the right to control their own likenesses, and their images should only be used in AI training data with their permission. Finally, each person should have the right to opt in or out of using this technology. With these safeguards in place, we feel that facial recognition software can be used in ways that are respectful, responsible, and beneficial.

Appendix A: Our Program (script.js)

Our program can be found here: https://marieke-thomas.github.io/facial-recognition-website/

In order to create a website that utilized the facial recognition algorithms we created with Google's Teachable Machine, we had to connect to their API in the script.js file twice. In order to connect to the API, we relied on sample code and documentation that enabled the user to connect a webcam and get usable data passed into the models. We modified this code to allow two webcams to be used at the same time. The first connection was in order to connect to our unbiased algorithm and the second was the more biased one. The `init()` and `init2()` functions (for biased and unbiased) are where we connect to the API, create the webcam and predictions on the website and get both to refresh continuously. The `loop()` and `loop2()` functions allow the program to continuously refresh the webcam and predict the gender at every new frame. Finally, the `predict()` and `predict2()` functions compare the frame to the model to determine the prediction that is displayed on the website.

At the bottom of our script.js file, starting on line 116 we have written lines of code that make buttons clickable. These buttons are in the Celebrity Example section of our website and allow the viewer to look closely at how famous celebrities are identified in a biased and unbiased model.

```
///////////////UNBIASED
    const URL2 =
"https://teachablemachine.withgoogle.com/models/2h73m7TG4/ ";

    let model2, webcam2, labelContainer2, maxPredictions2;

    // Load the image model and setup the webcam
    async function init2() {
        const modelURL2 = URL2 + "model.json";
        const metadataURL2 = URL2 + "metadata.json";

        // load the model and metadata
        // Refer to tmImage.loadFromFiles() in the API to support
files from a file picker
        // or files from your local hard drive
        // Note: the pose library adds "tmImage" object to your
window (window.tmImage)
        model2 = await tmImage.load(modelURL2, metadataURL2);
        maxPredictions2 = model2.getTotalClasses();

        // Convenience function to setup a webcam
        const flip2 = true; // whether to flip the webcam
        webcam2 = new tmImage.Webcam(200, 200, flip2); // width,
height, flip
        await webcam2.setup(); // request access to the webcam
        await webcam2.play(); //not currently playing on the
website
        window.requestAnimationFrame(loop2); //not currently
looping
        // console.log(webcam2)

        // append elements to the DOM

document.getElementById("webcam-container2").appendChild(webcam2.c
anvas);
```

```
        //
document.getElementById("webcam-container2").appendChild(document.
createElement("h2"));

        labelContainer2 =
document.getElementById("label-container2");
        for (let i = 0; i < maxPredictions2; i++) { // and class
labels

labelContainer2.appendChild(document.createElement("div"));
        }


    }

    async function loop2() {
        webcam2.update(); // update the webcam frame
        await predict2();
        window.requestAnimationFrame(loop2);
    }

    // run the webcam image through the image model
    async function predict2() {
        // predict can take in an image, video or canvas html
element
        const prediction2 = await model2.predict(webcam2.canvas);
        for (let i = 0; i < maxPredictions2; i++) {
            const classPrediction2 =
                prediction2[i].className + ": " +
prediction2[i].probability.toFixed(2);
            labelContainer2.childNodes[i].innerHTML =
classPrediction2;
        }
    }




//////////////BIASED

    // More API functions here:
    //
```

```
https://github.com/googlecreativelab/teachablemachine-community/tr
ee/master/libraries/image

    // the link to your model provided by Teachable Machine export
panel
    const URL =
"https://teachablemachine.withgoogle.com/models/P4ZhwqSMA/";

    let model, webcam, labelContainer, maxPredictions;

    // Load the image model and setup the webcam
    async function init() {
        const modelURL = URL + "model.json";
        const metadataURL = URL + "metadata.json";

        // load the model and metadata
        // Refer to tmImage.loadFromFiles() in the API to support
files from a file picker
        // or files from your local hard drive
        // Note: the pose library adds "tmImage" object to your
window (window.tmImage)
        model = await tmImage.load(modelURL, metadataURL);
        maxPredictions = model.getTotalClasses();

        // Convenience function to setup a webcam
        const flip = true; // whether to flip the webcam
        webcam = new tmImage.Webcam(200, 200, flip); // width,
height, flip
        await webcam.setup(); // request access to the webcam
        await webcam.play();
        window.requestAnimationFrame(loop);
          // console.log(webcam)


        // append elements to the DOM

document.getElementById("webcam-container").appendChild(webcam.can
vas);
        labelContainer =
document.getElementById("label-container");
        for (let i = 0; i < maxPredictions; i++) { // and class
labels
labelContainer.appendChild(document.createElement("div"));
        }
    }

    async function loop() {
        webcam.update(); // update the webcam frame
        await predict();
        window.requestAnimationFrame(loop);
    }

    // run the webcam image through the image model
    async function predict() {
        // predict can take in an image, video or canvas html
element
        const prediction = await model.predict(webcam.canvas);
        for (let i = 0; i < maxPredictions; i++) {
            const classPrediction =
                prediction[i].className + ": " +
prediction[i].probability.toFixed(2);
            labelContainer.childNodes[i].innerHTML =
classPrediction;
        }
    }

 // Below is code for the buttons in the Celebrity Test Cases
section of our website. When the button is pressed, it displays
the results for the corresponding celebrity with the algorithm
listed. This does not involve an API call; rather, we pretested
each celebrity picture against our two models and hard-coded their
results into the web page.

serenaButton = document.getElementById("serena-btn")
serenaResultsB = document.getElementById("serena-resultsB")
serenaButton.onclick = function(){
  serenaResultsB.innerHTML = "<p>Male: 100%</p><p>Female: 0%</p>"
};

serenaButton1 = document.getElementById("serena-btn1")
serenaResultsLB = document.getElementById("serena-resultsLB")
serenaButton1.onclick = function(){
  serenaResultsLB.innerHTML = "<p>Male: 0%</p><p>Female: 100%</p>"
```

```
};


jiminButton = document.getElementById("jimin-btn")
jiminResultsB = document.getElementById("jimin-resultsB")
jiminButton.onclick = function(){
  jiminResultsB.innerHTML = "<p>Male: 1%</p><p>Female: 99%</p>"
};

jiminButton1 = document.getElementById("jimin-btn1")
jiminResultsLB = document.getElementById("jimin-resultsLB")
jiminButton1.onclick = function(){
  jiminResultsLB.innerHTML = "<p>Male: 100%</p><p>Female: 0%</p>"
};

brunoButton = document.getElementById("bruno-btn")
brunoResultsB = document.getElementById("bruno-resultsB")
brunoButton.onclick = function(){
  brunoResultsB.innerHTML = "<p>Male: 17%</p><p>Female: 83%</p>"
};


brunoButton1 = document.getElementById("bruno-btn1")
brunoResultsLB = document.getElementById("bruno-resultsLB")
brunoButton1.onclick = function(){
  brunoResultsLB.innerHTML = "<p>Male: 100%</p><p>Female: 0%</p>"
};

kateButton = document.getElementById("kate-btn")
kateResultsB = document.getElementById("kate-resultsB")
kateButton.onclick = function(){
  kateResultsB.innerHTML = "<p>Male: 0%</p><p>Female: 100%</p>"
};

kateButton1 = document.getElementById("kate-btn1")
kateResultsLB = document.getElementById("kate-resultsLB")
kateButton1.onclick = function(){
  kateResultsLB.innerHTML = "<p>Male: 0%</p><p>Female: 100%</p>"
};
```

Appendix B: Explanation of Code Sorting the Photos

The faces that were used in this project were downloaded from the UTKFace Large Scale Face Database. The database includes a large number of files (over 20,000) which were downloaded in two separate folders labeled part1 and part2. Each photo is named with labels embedded, in the format: `[age]_[gender]_[race]_[date&time].jpg`

The gender is 0 for male, 1 for female, and the race is 0 for white, 1 for Black, 2 for Indian, 3 for Asian, and 4 for other including Latino. Below is the code that we used to sort this large batch of photos, along with a brief explanation.

```
Mariekes-Air:downloads marieke$ mkdir facial_recognition_pics
Mariekes-Air:downloads marieke$ mv part1 facial_recognition_pics/
Mariekes-Air:downloads marieke$ mv part2 facial_recognition_pics/
Mariekes-Air:downloads marieke$ cd facial_recognition_pics/
Mariekes-Air:facial_recognition_pics marieke$ mkdir all_pics
Mariekes-Air:facial_recognition_pics marieke$ mv part1/* all_pics
-bash: /bin/mv: Argument list too long
Mariekes-Air:facial_recognition_pics marieke$ find part1 -name
"*.jpg" -exec mv {} all_pics \;
Mariekes-Air:facial_recognition_pics marieke$ find part2 -name
"*.jpg" -exec mv {} all_pics \;
```

First, we wanted to combine all of the photos into one directory with the location downloads/facial_recognition_pics/all_pics. We made the facial_recognition_pics directory and moved the part1 and part2 directories, containing the photos, into it. Then we created an all_pics directory. We first attempted to simply move all of the photos from part1 into all_pics. However, this yielded an error stating that the Argument list was too long. Essentially there were too many photos to be moved using the mv command. In order to move all of the photos, we first had to find them by going into the part1 folder and finding every file that ended in ".jpg". We then used the -exec to execute code for all the images that fall under this category with the mv command. This moved every photo from part1 into the all_pics directory. We did the same with part2.

```
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9]_0_0_* white_men
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9][0-9]_0_0_* white_men
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9][0-9][0-9]_0_0_* white_men
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9]_1_0_* white_women
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9][0-9]_1_0_* white_women
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9][0-9][0-9]_1_0_* white_women
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9]_0_* all_men
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9][0-9]_0_* all_men
-bash: /bin/cp: Argument list too long
Mariekes-MacBook-Air:facial_recognition_pics marieke$ find
all_pics -name "[0-9][0-9]_0_*" -exec cp {} all_men \;
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9][0-9][0-9]_0_* all_men
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9][0-9][0-9]_1_* all_women
Mariekes-MacBook-Air:facial_recognition_pics marieke$ cp
all_pics/[0-9]_1_* all_women
Mariekes-MacBook-Air:facial_recognition_pics marieke$ find
all_pics -name "[0-9][0-9]_1_*" -exec cp {} all_women \;
```

Above is the code that we used to make the folders of training data for our machine learning algorithm. This code was executed in Bash in the MacBook terminal. Bash does not use Regular Expressions, so instead the files had to be sorted using glob pattern matching, which is more limited. The basic pattern for white men would be age_0_0_longNumber. However, the age can be one, two, or three digits. In regular expressions we could use [0-9]+ to account for the different numbers of digits, but this is not possible in glob, so we used 3 different lines of code to copy and move the photos of white men of different ages. We similarly moved white women using 3 lines of code to account for ages with one, two, and three digits. When we tried to move all men with 2-digit ages and all women with 2-digit ages, we were unable to use "cp" as the number of files was too large. As before, we used the find function to execute code over all files that matched our criteria. For example, to move all men with 2-digit ages, we used: `find all_pics -name "[0-9][0-9]_0_*" -exec cp {} all_men \;`.

Works Cited:

Alms, N. (2022, May 11). ICE has assembled a "surveillance dragnet" with facial recognition and data, report says. *FCW*. https://fcw.com/digital-government/2022/05/ice-has-assembled-surveillance-dragnet-facial-recognition-and-data-report-says/366822/

Apple. (2022, March 14). *Use Face ID on your iPhone or iPad Pro*. Apple Support. https://support.apple.com/en-us/HT208109

Biometric Information Privacy Act (BIPA). (2022, April 29). ACLU of Illinois. https://www.aclu-il.org/en/campaigns/biometric-information-privacy-act-bipa

Bohannon, J. (2015, December 21). Eyewitness testimony may only be credible under these circumstances. *Science*. https://www.science.org/content/article/eyewitness-testimony-may-only-be-credible-under-these-circumstances

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Conference on Fairness, Accountability and Transparency*, 77–91.

Chowdhury, M. (2022, July 6). Is Facial Recognition Really Accurate? *Analytics Insight*. https://www.analyticsinsight.net/is-facial-recognition-really-accurate/

Cuthbertson, A. (2019, August 1). Hong Kong protesters use lasers to avoid facial recognition cameras and blind police. *The Independent*. https://www.independent.co.uk/tech/hong-kong-protests-lasers-facial-recognition-ai-china-police-a9033046.html

Facial Recognition - NYPD. (n.d.). https://www.nyc.gov/site/nypd/about/about-nypd/equipment-tech/facial-recognition.page

Freeman, A., & Delger, D. (2020, May 1). Bite Mark Evidence: What You Should Know About This Debunked Science. *Innocence Project*. https://innocenceproject.org/what-is-bite-mark-evidence-forensic-science/

Teachable Machine. (n.d.). Google. https://teachablemachine.withgoogle.com/

Gupta, V., & Mallick, S. (2022, November 7). *Face Recognition : A 30000 feet view | LearnOpenCV*. LearnOpenCV – Learn OpenCV, PyTorch, Keras, Tensorflow With Examples and Tutorials. https://learnopencv.com/face-recognition-an-introduction-for-beginners/

Hartsock, K., & Woods, K. (2018, August 9). *Using Spotlight to investigate human trafficking*. Thorn. https://www.thorn.org/blog/using-spotlight-to-investigate-human-trafficking/

Johnson, K. (2022, March 7). How Wrongful Arrests Based on AI Derailed 3 Men's Lives. *WIRED*. https://www.wired.com/story/wrongful-arrests-ai-derailed-3-mens-lives/

Keyes, O. N. S. (2019, March 18). The Government Is Using the Most Vulnerable People to Test Facial Recognition Software. *Slate Magazine*. https://slate.com/technology/2019/03/facial-recognition-nist-verification-testing-data-sets-children-immigrants-consent.html

Law Policy. (2008). Strengthening forensic science in the United States : a path forward. *National Academies Press EBooks*.

Nessen, S. (2021, March 31). MTA's Broad Surveillance Camera Expansion Did Little To Reduce Crime. *The Gothamist*. https://gothamist.com/news/mtas-broad-surveillance-camera-expansion-did-little-to-reduce-crime

Ng, A. (2020, August 11). How China uses facial recognition to control human behavior. *CNET*. https://www.cnet.com/news/politics/in-china-facial-recognition-public-shaming-and-control-go-hand-in-hand/

Noorden, R. V. (2020, November 18). The ethical questions that haunt facial-recognition research. *Nature*.

https://www.nature.com/articles/d41586-020-03187-3?error=cookies_not_supported&code=2cafa1b0-449c-4a29-bdfe-f6df551cb337

Pesenti, I. A. O. J. V. P. (2021, November 3). An Update On Our Use of Face Recognition. *Meta*. https://about.fb.com/news/2021/11/update-on-use-of-face-recognition/

Phillips, P. J., Jiang, F., Narvekar, A., Ayyad, J., & O'Toole, A. J. (2011). An other-race effect for face recognition algorithms. *ACM Transactions on Applied Perception*, *8*(2), 1–11. https://doi.org/10.1145/1870076.1870082

Priks, M. (2015). The Effects of Surveillance Cameras on Crime: Evidence from the Stockholm Subway. *The Economic Journal*, *125*(588), F289–F305. https://doi.org/10.1111/ecoj.12327

Privacy Principles for Facial Recognition Technology in Commercial Applications. (2019). Future of Privacy Forum. https://fpf.org/wp-content/uploads/2019/03/Final-Privacy-Principles-Edits-1.pdf

Simonite, T. (2019, June 19). How Facial Recognition Is Fighting Child Sex Trafficking. *WIRED*. https://www.wired.com/story/how-facial-recognition-fighting-child-sex-trafficking/

Song, Y., & Zhang, Z. (n.d.). *UTKFace*. UTKFace: Large Scale Face Dataset. https://susanqq.github.io/UTKFace/

Thorn. (2022, August 30). *Spotlight: Human Trafficking Intelligence and Leads*. https://www.thorn.org/spotlight/

What is Facial Recognition - Beginner's Guide to Face Analyzer Software and Machine Learning - AWS. (n.d.). Amazon Web Services, Inc. https://aws.amazon.com/what-is/facial-recognition/

(2022, February 2). Beneath the Statistics: The Structural and Systemic Causes of Our Wrongful Conviction Problem. *Georgia Innocence Project*. https://www.georgiainnocenceproject.org/2022/02/01/beneath-the-statistics-the-structural-and-systemic-causes-of-our-wrongful-conviction-problem/

5 most common uses of facial recognition. (2022, January 25). *NEC*. https://www.nec.co.nz/market-leadership/publications-media/5-most-common-uses-of-facial-recognition/