**Facial Recognition: Distinguishing Masked from Unmasked Images**

Zach Quinn

Winter 2021

GitHub: <https://github.com/Zachlq/Professional_Portfolio/tree/main/TensorFlow%20Mask%20Detection>

Bellevue University

**Introduction**

With facial coverings, including masks, becoming a necessary staple of culture within the past year, image classification algorithms are facing a daunting challenge: Obscured facial features. Facial recognition is an emerging subfield within artificial intelligence and data science. Commercial and public spaces leverage facial recognition to preserve security and even to reveal consumer insights. The application for facial recognition ranges from micro transactions like unlocking one’s phone to the identification of protesters. Government watchdog group the American Civil Liberties Union estimates that millions of Americans have had their images captured and stored in global databases. Specifically, the ACLU cites a report that suggests the FBI has access to over 640 million facial images, a number that is nearly double the population of the United States of America (Guliani, 2019). The revelations from the FBI suggest that, controversial or not, mass facial recognition techniques are being deployed and refined for purposes such as predictive policing, fraud detection and AI-powered national security initiatives (Guliani, 2019). Law enforcement’s leveraging of facial recognition to identify protesters in the many riots that occurred within the past five years has resulted in technology giants like Microsoft pledging not to sell facial recognition applications to law enforcement departments that may be conducting operations with malice or other apparent prejudice. However, with the introduction of the coronavirus to the world in 2019, these rather conventional classification algorithms are encountering difficulties when comparing images of masked to unmasked faces. In an attempt to circumvent this latest obstacle to facial recognition, this project will construct and deploy a multi-layered neural network to facilitate the classification of a masked and unmasked face dataset.

**Business Problem**

The facial recognition market is currently valued at approximately 3.8 billion dollars, with analysts projecting that figure will increase to nearly 9 billion dollars by 2025 and nearly 15 billion dollars by 2027 (Fortune Business Insights, 2020). Representing one of the fastest growing fields within AI, facial recognition, whether consumers are aware or not, has become engrained in western culture. In 2017 Apple launched the iPhoneX and, with it, debuted Face ID, a feature that allows users to unlock their phones simply by allowing the device to capture and verify a profile image (Heilweil, 2020). In order to function, the camera creates a 30,000-dot dimensional map of a subject’s facial features (Apple, 2020). Both governmental and corporate entities have refined and deployed facial recognition algorithms and, chances are, if you’ve been in public within the last five years, an image of your face has been captured and stored within a massive database of images. However, with the introduction of masks into society within the past year, these once-precise image classification algorithms are having difficulty achieving benchmark accuracy scores, with some programs unable to perform above fifty percent (Telford, 2020).

The National Institute for Standards and Technology revealed that variables such as mask color, size and shape can impede the functionality of classification systems (NIST, 2020). Already, opponents of facial recognition contend that these systems fail to correctly identify certain skin tones and even have difficulty distinguishing between male and female subjects (Simonite, 2020). With the onset of pandemic-related mask mandates occurring less than a year ago, one of the most significant challenges to mask facial recognition is the lack of available images to train facial recognition algorithms (Simonite, 2020). Since numeric and text-driven algorithms typically require large datasets (millions of data points) to train a model to a high degree of precision, experts assert that, to achieve any degree of accuracy with facial recognition, models will require at minimum, hundreds of thousands of images of masked faces (Simonite, 2020). Consequently, contemporary facial recognition algorithms face a dual problem: The obscuring of features such as the human nose and mouth typically used to generate data points and a lack of available masked face images (Simonite, 2020). Therefore, it is necessary to discover or mine a large amount of image data to ensure that a model is sufficiently trained on a robust dataset, and tested with images that contain disparities such as differing colors and shapes, to comprehensively assess an ML model exposed to a dataset of masked images.

**The Project**

This project seeks to develop and deploy a multi-layered neural network in order to successfully distinguish between a dataset of masked and unmasked facial images. Given the data constraints outlined in the previous portions of this essay, this project endeavors to achieve a high level of precision on a small to medium-sized dataset with the hope of the model being applied to a larger validation set when such a variety of images is made publicly available. Instead of converting pixels to data points and developing a multi-dimensional map of a subject’s face, the project intends to approach this problem as a classification issue rather than a real-time implementation of facial identification. Any datasets used to both train and validate a machine learning model will include images of masked and unmasked faces with varying features such as different color masks, different size masks and masks worn in various fashions to obscure different portions of the face. The goal is that a model can be trained to recognize, in a broader sense, the presence of a mask on a human face without being misled by the variables outlined in the introduction.

**Hypothesis**

This project aims to leverage existing deep learning libraries and image classification algorithms to facilitate the automated distinction between images of masked and unmasked images. This project aims to train a convolution neural network (CNN) or recurrent neural network (RNN) to discern whether or not an individual’s face is sufficiently covered according to comply with Centers for Disease Control (CDC), World Health Organization (WHO) and White House airborne virus mitigation guidelines. Therefore, the hypothesis for this experiment is as follows: By training a neural network on a sufficiently variable dataset of unmasked and masked facial images, it will be possible to train a machine learning model to correctly distinguish between images of both masked and unmasked faces with a precision rate greater than 70%.

**10 Research Questions**

1. What epoch or batch size is most impactful on the overall functionality of a CNN trained to classify masked or unmasked images?
2. How might an improperly worn mask impact the precision of the model?
3. How might photo resolution or other image quality concerns impact the performance of the model?
4. In what ways can data augmentation be performed to increase the size and variety of data?
5. What utility might a bounding box serve in determining and illustrating spatial relationships between a mask and face when it comes to classification?
6. What visualizations can we conceive for image-based data?
7. How might we train a similar model to determine whether or not an individual in a photo is following the six feet distancing guideline?
8. How might a more equitable train/test split (i.e. 70-30 or 60-40) impact the performance of the model?
9. How might this classification model fare on non-medical masks? Would it be the same approach?
10. What other neural networks might be able to achieve the same degree of precision? For instance, why might a CNN be more effective than an RNN? In what ways can the performance of such algorithms be evaluated?

**Data**

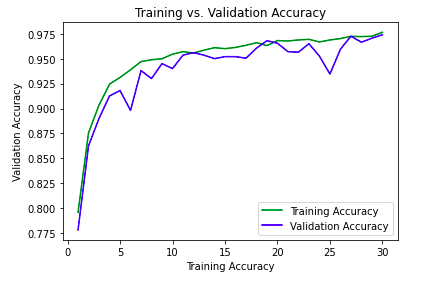
The image dataset is small by professional standards, comprising only approximately 10000 images. The dataset was derived from The Assembly, a Dubai-based academic think tank headed by luminaries like the head of multinational telecommunications firm Qualcomm, which partitioned the raw data into two classes: ‘with’ and ‘without.’ These designations, predictably, indicate that the individuals in the photos are either wearing or not wearing a mask. In addition to ‘normal’ individuals, the dataset also contains unmasked and masked images of celebrities and notable public figures.

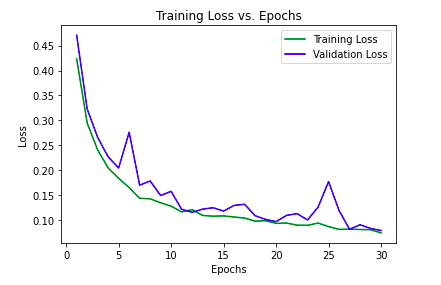
**Methodology**

In order to make a binary prediction regarding whether or not an individual is wearing a mask in a photo, this project utilized a multi-layer convolution neural network (CNN).  
In order to avoid consuming too much disk space with thousands of images, the data was uploaded to Google Collaboratory and unzipped in an IDE. The data was assigned to a directory variable to facilitate a smooth pipeline to an image generator. Since the dataset was medium to small (10,000 images), data augmentation was performed to diversify the kinds of data both the training and validation models were exposed to. Keras’ image data generator was used to perform the augmentation, which included horizontal and vertical flips as well as a shuffle. The training and validation sets were split with an 80 - 20 partition. With the data cleaned and processed, the model was created by first initiating a Sequential object. Conv2D, Dense and Maxpool layers were added along with relu activations for the top layers and a sigmoid activation for the bottom layer. The chosen number of epochs was 30 while the batch size was 8. 30 was chosen to allow the model sufficient time to iterate and train to reach high accuracy and low loss rates. The small batch size was selected to be proportional to the relatively small dataset. The training and validation loss as well as training and validation accuracy were plotted on a line graph to evaluate and visualize the model’s performance.

**Results**

Below are graphs detailing the accuracy and loss of the model on the training and validation sets. The data inserted below represents the first time running the model in its entirety. Therefore, these results are subject to change as epoch, batch size and other parameters are tuned for the final iteration of this project.



****

The performance of neural networks is typically evaluated by two metrics: Loss and accuracy. The loss of a model is a penalty imposed on the model for incorrect guesses. In other words, loss can be thought of as a model’s error rate. Accuracy is the percentage of correct predictions on the validation, or test, set. On this small dataset, the model was able to achieve an accuracy rate of 97 percent, exceeding the project’s goal of an accuracy rate above 85 percent. Additionally, the loss was fairly minimal, with the highest rate of loss, 50 percent, occurring in the first epoch. This was predictable though because the model had not yet been exposed to all of the data. By the 30th epoch the loss rate had fallen to 7 percent.

**Discussion**

While this project may look good on paper, facial recognition is still a relatively new and complex science. One of the field’s most significant problems is the lack of abundant and diverse databases. This will be a difficult challenge to address so long as the public perceives facial recognition to be both an invasion of privacy and even a discriminatory activity. However, this is a vicious cycle because facial recognition technology cannot improve until it is exposed to datasets that are more representative of the population at large. The consequences of bias in facial recognition algorithms are impacting businesses and individuals. Notably, New York City’s Mass Transit Association cut funding to a facial recognition pilot after it performed disastrously, achieving a 100 percent error rate (Kraus, 2021). More research and news coverage is emerging to support the assertion that facial recognition technology is inherently biased. Such algorithms are less likely to be able to identify non-white individuals. This bias has already resulted in the arrests of three black men who were mistakenly identified by the technology (Kraus, 2021).

**Conclusion**

A convolutional neural net is an efficient and precise means of conducting a basic classification of masked and unmasked individuals in photos. Even as authorities begin to regain control of the pandemic and individuals reclaim lost personal freedoms, data analytics firms and security agencies are increasingly interested in how to train algorithms to conduct facial recognition tasks of masked individuals. As vaccine distribution accelerates in the coming months, masks will undoubtedly be phased out of widespread use. However, the apparent threat of illness will likely usher in a new age in which masks become optional in American society during surges of virus cases within certain communities or seasons, such as flu season. The controversy of mask mandates has collided with the privacy concerns associated with purportedly invasive facial recognition techniques. In the future, it will be necessary to design and implement image classification algorithms that take into account the possibility that an individual will obscure a portion of their face for health or security reasons. The CNN implementation proposed in this paper provides a simple, efficient and reproducible technique for achieving precise and versatile facial screenings on masked and unmasked subjects.

**References**

Apple. (2020). About Face ID Technology. *Apple*. <https://support.apple.com/en->us/HT208108#:~:text=The%20technology%20that%20enables%20Face,infrared%20image%20of%20your%20face.

Fortune Business Insights. (2020). Facial Recognition Market Size, Share and COVID-19 Impact Analysis. *Fortune Business Insights*. <https://www.fortunebusinessinsights.com/industry-reports/facial-recognition-market-101061>

Guliani, N. (2019). The FBI Has Access to Over 640 Million Photos of Us Through Its Facial Recognition Database. *American Civil Liberties Union*. <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/fbi-has-access-over-640-million-photos-us-through>

Heilweil, R. (2020). How can we ban facial recognition when it’s already everywhere? *Vox*. [https://www.vox.com/recode/2020/7/3/21307873/facial-recognition-ban-law-enforcement-apple-google-facebook](%09https://www.vox.com/recode/2020/7/3/21307873/facial-recognition-ban-law-enforcement-apple-google-facebook)

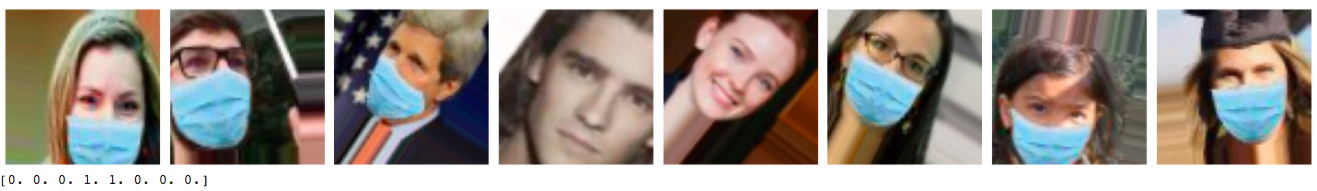
Kraus, R. (2021). 9 Scary Revelations from 40 Years of Facial Recognition Research. Mashable. <https://mashable.com/article/facial-recognition-databases-privacy-study/>

NIST. (July 2020). NIST Launches Studies into Masks’ Effect on Face Recognition Software. *National Institute of Standards and Technology*. <https://www.nist.gov/news-events/news/2020/07/nist-launches-studies-masks-effect-face-recognition-software>

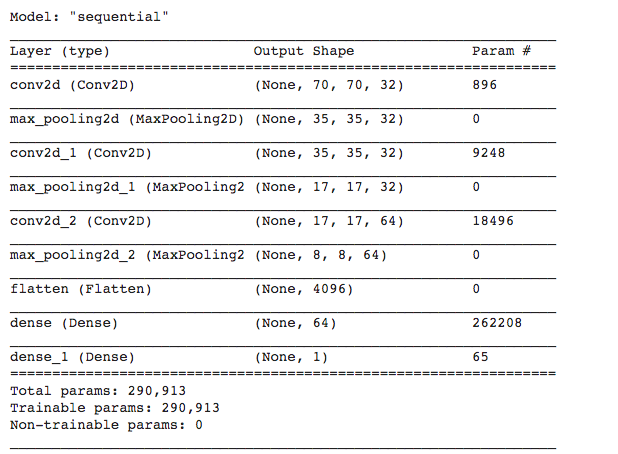
Simonite, T. (2020). How Well Can Algorithms Recognize Your Masked Face? *Wired*. <https://www.wired.com/story/algorithms-recognize-masked-face/>

Telford, T. (2020). Top facial recognition tech is thrown off by masks, study says, but that could soon change. *The Washington Post*. <https://www.washingtonpost.com/business/2020/07/29/masks-facial-recognition-error-rates/>

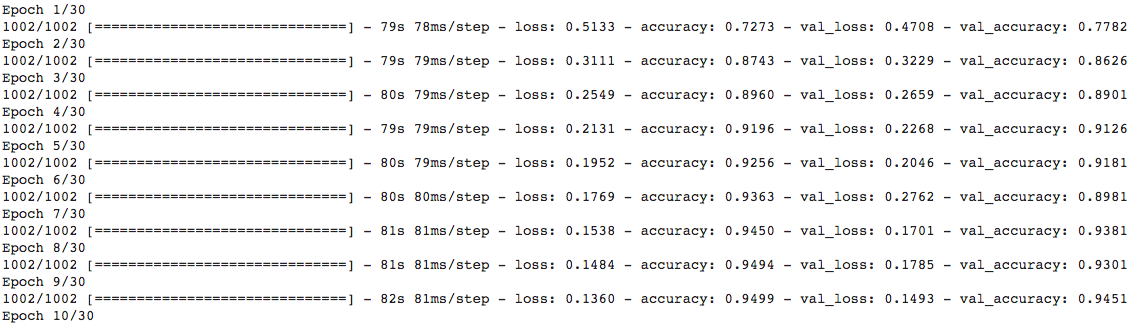
**Appendix A: Dataset (Selected Images)**



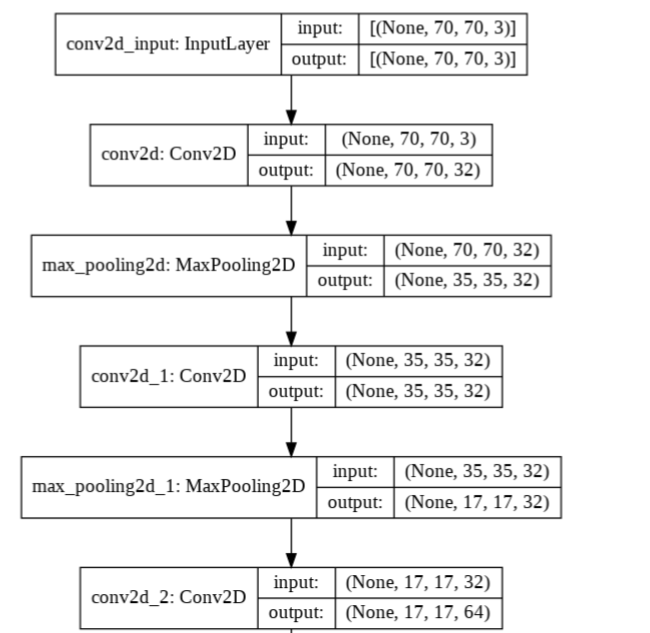
**Appendix B: Model Summary**



**Appendix C: Performance of First Ten Epochs**

****

**Appendix D: Model Illustration**



**Appendix E: Data Repository**

A zipped file of 10,000+ images, courtesy of the Assembly

<https://github.com/The-Assembly/Build-A-Face-Mask-Detector-With-TensorFlow/blob/main/FaceMask%20detection/data.zip>