Enhanced Facial Detection: Age and Mask Identification for Diverse Applications

Omar Ali
Department of Computer Science
California State University, Fresno
Fresno, California
omarali112511@mail.fresnostate.edu

Ulysses Ochoa
Department of Computer Science
California State University, Fresno
Fresno, California
ulysses_ochoa@mail.fresnostate.edu

Nicholas Rodriguez
Department of Computer Science
California State University, Fresno
Fresno, California
nickrodz@mail.fresnostate.edu

ABSTRACT

The ability to detect and analyze facial features has become an essential technology across various fields including security systems, health monitoring, and personalized user experiences. This paper introduces a sophisticated deep learning methodology that employs Convolutional Neural Networks (CNNs) to determine age and detect mask-wearing status from facial images. We developed two specialized neural network models, meticulously trained on expansive and diverse datasets: the UTKcropped dataset for age detection and the Face Mask Detection Mask Dataset for mask identification.

These models are designed to decode both high-level and subtle facial features, thereby enabling accurate age categorization and reliable mask detection. To optimize performance, we thoroughly explored different neural network architectures and fine-tuned hyperparameters, aiming to enhance both the accuracy and efficiency of the models.

Preliminary results are promising, demonstrating the models' ability to accurately estimate age with a validation accuracy of 62.39%, and to detect various types of masks with a high validation accuracy of 94.25%. The applications for these models are broad, ranging from improving security measures by integrating age and mask detection to supporting public health efforts through the automated monitoring of mask usage in real-time environments. Additionally, we have created a website to demonstrate the practical capabilities of these models, offering users a hands-on experience with the technology.

1 INTRODUCTION

Facial feature detection has significantly evolved from traditional techniques, which often relied on manual feature extraction and simple pattern recognition, to more sophisticated methodologies utilizing deep learning. This shift is particularly evident in applications requiring the real-time identification of specific facial attributes, such as age and mask-wearing status. Our project capitalizes on this technological evolution by deploying CNNs to analyze and interpret facial data with remarkable accuracy.

In this venture, we specifically target two pertinent aspects of facial detection: age estimation and mask detection. Given the global emphasis on health monitoring and enhanced security systems, these features are not only relevant but critical in managing public health events and security protocols. For instance, age detection can significantly augment demographic analytics without intruding on personal privacy, while mask detection has become crucial in managing health protocols during events such as the COVID-19 global pandemic.

To address these needs, we developed two separate models: one for age detection and another for mask detection. Each model is designed to process facial data effectively, ensuring accurate and timely results that are crucial for real-time applications.

The functionality of these models is showcased through a website that acts as a practical demonstration platform. This website integrates the models into a real-world application by employing live video feeds from a device's camera. As the video captures faces, our CNNs analyze the data in real-time, estimating age and identifying whether individuals are wearing masks. This setup not only tests the dependability of our models under realistic conditions but also provides an immediate, practical application that could be scaled and amplified for wider use in various sectors.

2 RELATED WORKS

2.1 Research Papers

Chang Kong et al.'s study, "RSFAD: A Large-Scale Real Scenario Face Age Dataset in the wild," [1] introduces a groundbreaking dataset aimed at refining age estimation technologies under uncontrolled environmental conditions. The RSFAD dataset includes over 85,000 images collected from surveillance cameras, providing a realistic basis for developing and testing age estimation algorithms. The study not only addresses the common biases in age and gender distribution found in existing datasets but also integrates mask-wearing variables, reflecting the need for adaptable technologies in times of health emergencies. By using this dataset, the researchers demonstrate improvements in age estimation accuracy, showing the critical role of diverse and realistic datasets in training more effective and applicable models.

Rasha Ragheb Atallah et al.'s review, "Face Recognition and Age Estimation Implications of Changes in Facial Features," [2] provides a thorough analysis of the advancements and challenges in face recognition technologies. This critical survey highlights the impact of aging on facial features and discusses the effectiveness of various methodologies and technologies within the field. The study meticulously examines over 72 articles to shed light on the evolution of recognition techniques, emphasizing the necessity for robust models capable of accurately predicting age and identifying faces across a spectrum of conditions and demographics. By addressing significant issues such as the variability in age-related facial characteristics and the limitations imposed by low-resolution images and diverse facial expressions, this paper illuminates the critical need for enhancements in current systems to improve real-world application accuracy.

2.2 Datasets

UTK Face Cropped Dataset: This Kaggle dataset comprises 23,709 well-cropped and aligned face images that span a broad age range from 0 to 116 years. It is utilized to develop the age category classification model, providing a diverse basis for understanding and predicting age-related facial features across a wide demographic.

Face Mask Detection Dataset: Also sourced from Kaggle, this dataset is employed for the mask binary identification model and contains a total of 11,792 face images categorized according to mask usage. The dataset is meticulously organized into several subsets: two folders, each containing 400 images, are used for validation of mask and no-mask scenarios, summing up to 800 images. For training, two additional folders comprise 5,000 images each for masked and unmasked faces. The testing set includes 483 images with masks and 509 images without, allowing for comprehensive evaluation of the mask detection model under varied real-world conditions.

2.3 Open-Source Programs and Resources

For the development of the website demonstration associated with the facial recognition models, the project utilized methodologies and code examples from the OpenCV Library. Specifically, the guide "Face Detection using Haar Cascades," available through the official OpenCV documentation, was instrumental.

3 METHODS

The development of the facial detection project commenced with the creation of two distinct models: one for age estimation and another for mask detection. These models were trained using the UTKcropped Face dataset and the Face Mask Detection dataset, adhering to an 80:20 training-to-validation ratio for the UTKcropped dataset and a 25:2 ratio for the Face Mask Detection dataset, respectively.

For both age detection and mask identification, Convolutional Neural Networks (CNNs) were employed, utilizing essential libraries such as Keras and NumPy to optimize computational efficiency. These models were architecturally similar, with the key difference lying in their classification tasks. The age estimation model used a categorical classification approach to distinguish among various age ranges, requiring it to handle multiple categories effectively.

First, we normalized the input data and employed several convolutional layers with batch normalization and ReLU activation to extract meaningful features. Residual connections ensure efficient training in deeper layers. The model also uses separable convolutions to reduce computational load and maximize feature extraction. A final Global Average Pooling layer condenses these features before passing them to a dropout layer, preventing overfitting. The dense output layer with a sigmoid activation function provides a binary classification for mask detection. The dense output layer with a softmax activation function provides an 8-class categorical classification for the age classification. The Adam optimizer is utilized for its adaptive learning rate, which efficiently minimizes binary cross-entropy loss for the mask model and categorical cross-entropy loss for the age model while improving overall accuracy.

The mask detection model, in contrast, was designed for binary classification, tasked with determining whether an image depicted a masked or unmasked face. Despite this difference in classification complexity, both models underwent a similar training process. Parameters were iteratively adjusted to refine the models' abilities to classify images accurately. To ensure continuity and progress, the state of each model was saved to a designated folder after each training iteration.

Post-training, the effectiveness of both models was evaluated through a series of rigorous tests. Python scripts were employed to perform hard predictions on designated folders containing images from the training datasets. The age detection model was tasked with classifying images into specific age ranges, aiming for high accuracy within each category. Similarly, the mask detection model was tested to differentiate images based on mask usage effectively.

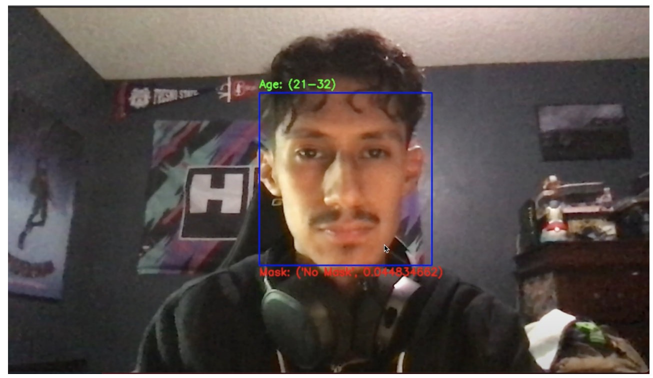

Figure 1: A screenshot of the demonstration website showing face detection in action. The interface displays a live video feed with a detection box around the user's face, indicating the estimated age range and mask status.

Upon successful validation, both models were integrated into a demonstration website. This platform utilized live camera feeds to assess and display the user's age range and mask status in real-time. The website was crafted using HTML and CSS for the front end, while PHP managed server-side requests. These requests activated Python scripts that processed live feeds, employing both models for real-time predictions. The OpenCV library, with its robust Haar Cascade Classification algorithm, was instrumental in facilitating real-time face detection, essential for the dynamic functioning of our models on the website.

4 RESULTS

For the Age Model, the training and validation results were analyzed using graphs plotted in Google Sheets. The accuracy and loss graphs illustrate the trends over 15 epochs. The training and validation accuracies for the Age Model were relatively stable, though fluctuations were noticeable, particularly in validation accuracy.

The Age Model was divided into eight categories labeled 'Test Category' in the table displaying the age estimation results. Category 1 represents the youngest age group, with each subsequent category encompassing progressively older age ranges up to the highest age group. For each test category, the table indicates how accurately the model classified images within that category into the

Test Category	0-3	4-7	8-14	15-20	21-32	33-43	44-53	54-100	Total # of Images
1	64.39%	0.00%	0.00%	0.00%	1.52%	6.06%	0.00%	28.03%	396
2	34.01%	3.40%	4.08%	0.00%	4.08%	35.37%	0.00%	19.05%	147
3	7.66%	0.90%	16.22%	0.00%	8.11%	49.10%	0.00%	18.02%	222
4	0.84%	0.00%	5.04%	1.26%	32.77%	52.94%	0.00%	7.14%	238
5	0.28%	0.00%	0.06%	0.00%	19.64%	66.44%	0.00%	13.58%	242
6	0.00%	0.00%	0.00%	0.00%	7.33%	63.75%	0.00%	28.92%	778
7	0.00%	0.00%	0.00%	0.00%	1.68%	43.65%	0.24%	54.44%	417
8	0.00%	0.00%	0.00%	0.00%	1.68%	13.53%	0.13%	84.76%	761

Figure 2: A table displaying the age group classifications, accuracies, and the number of images per group used in the facial detection project. Each column represents a different age group, listing the corresponding accuracy percentage and the total number of images classified within that group.

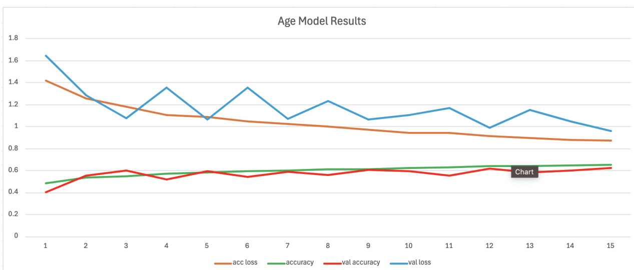

Figure 3: A line graph displaying the age model results over 15 epochs, showing training and validation accuracy (green and orange lines) and loss (blue and red lines) respectively.

corresponding age ranges. A higher match between the predicted age group and the test category indicates better accuracy.

For instance, in Test Category 6, which represents the age range 33-43, the model classified images within this group with an accuracy of 63.75%. Similarly, in Test Category 8, representing ages 54-100, the model achieved an accuracy of 84.76%. Overall, the model demonstrated higher accuracy in detecting babies and seniors. However, for younger adults in the age range 21-32 (Test Category 5), it often confused them with adults in the age group 33-43 (Test Category 6).

The Mask Model's training and validation accuracy and loss were graphed using Python's Matplotlib library. The training and validation accuracy vs. loss graphs provide valuable insights into the model's learning process. Initially, we ran the model for only two epochs, but we extended the number to six to allow for a longer training period, enabling the model to generalize better. The Mask Model classified masked and unmasked faces with high accuracy. For Test Category 0 (Mask), the model achieved a near-perfect accuracy of 99.5%, while for Test Category 1 (No Mask), it accurately classified 93% of the images. These results were based on the validation set of images.

5 CONCLUSION

Our project is centered on achieving high accuracy in age determination, both with and without the presence of a mask. The

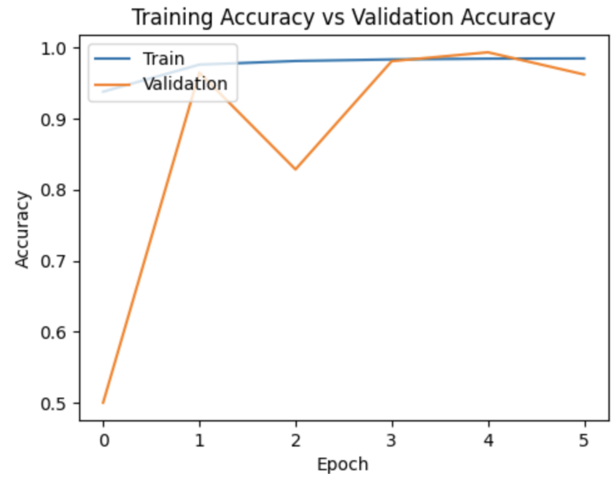

Figure 4: A graph showing training and validation accuracy over five epochs for the mask identification model, with training in blue and validation in orange.

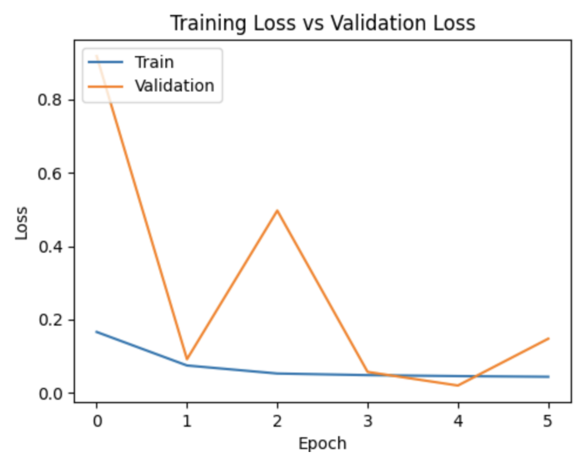

Figure 5: A graph depicting training and validation loss over five epochs for the mask identification model, with training in blue and validation in orange.

Test Category	Mask	No Mask	Total # images
0 (Mask)	99.50%	0.50%	400
1 (Mask)	7%	93%	400

Figure 6: A table summarizing the mask detection accuracy for images with masks (Category 0) and without masks (Category 1), showing percentages of correct classifications alongside the total count of 400 images tested per category.

age model attained a training accuracy of 65.24% and a validation accuracy of 62.39%, indicating satisfactory performance with room for enhancement. Conversely, the mask model exhibited superior accuracy, achieving a training accuracy of 94.97% and a validation accuracy of 94.25%. However, challenges arose when applying these models to new images, resulting in a notable decrease in accuracy, likely attributable to overfitting. While the models demonstrate proficiency with training data and validation sets resembling the training data, they encounter difficulty with unseen data. Notably, our models use RGB in detecting the colorization of images which poses an issue as the website utilizes grayscale for face detection, leading to inaccuracies in age estimation and mask detection. A possible improvement we could make in the future is finding an alternative to the Haar Cascade face detector of the website where it takes input of RGB instead of grayscale. More testing can be done with the models, specifically the age model to improve the accuracy with a wider range of age categories.

ACKNOWLEDGMENTS

Our group is immensely grateful for the support and guidance provided by Professor Amith Kamath Belman throughout the course of our project. We also wish to acknowledge the comprehensive and insightful content of CSci 158: Applied Biometric Security, which greatly enhanced our understanding and application of the subject matter. Additionally, we thank the Department of Computer Science at California State University, Fresno, for fostering an environment conducive to our academic and practical pursuits in this field.

REFERENCES

[1] Qiuming Luo Chang Kong and Guoliang Chen. 2021. RSFAD: A Large-Scale Real Scenario Face Age Dataset in the wild. 15-18 (Dec. 2021), 7. https://doi.org/10.1109/FG52635.2021.9667040

[2] Maizatul Akmar Ismail Sherin Ali Abdelrahman Rasha Ragheb Atallah, Amirrudin Kamsin and Saber Zerdoumi. 2018. Face Recognition and Age Estimation Implications of Changes in Facial Features: A Critical Review Study. 18 (May 2018), 15. https://doi.org/10.1109/ACCESS.2018.2836924