Grenze International Journal of Engineering and Technology, June Issue

Real Time Facial Mask Detection and Gender Identification

D. Shivaram[1], D. Sreeram[2], Joshila Grace L.K[3], P. Asha[4], Jancy[5] and Mercy Paul Selvan[6]
[1-6]Sathyabama Institute of Science and Technology, Chennai – 600119, Tamilnadu, India
Email: joshilagracejebin@gmail.com

***Abstract*— This work presents a real time system that focuses on facial detection, mask detection and gender identification. The proposed framework uses advanced computer vision techniques to capture video input, classify faces using pre trained Haar cascades or deep learning models, classify on mask status and gender classification using CNNs. With video frames being processed efficiently by the system, its public safety and demographic analysis benefits can be immediate. This integration of making facial recognition work with mask compliance monitoring and gender classification is a powerful indication of the use of artificial intelligence in public health, security, and analytics scenarios. The model is trained on a dataset with images of people with and without masks, detecting the presence of a mask in real-time for applications Gender identification is the task of predicting a person's gender based on their facial features. The system predicts whether the person is male or female based on visual cues from their face, commonly used in marketing, security, and personalized systems. Both tasks use deep learning to analyze facial features efficiently. The future will see emotion recognition, multi-threading for speed, and edge computing for running on resource constrained devices.**

***Index Terms*— Real-time facial detection, mask detection, gender classification, convolutional neural networks (CNNs), computer vision, public health safety, demographic analysis, artificial intelligence (AI).**

I. INTRODUCTION

Artificial intelligence (AI) and computer vision technologies have revolutionised how humans interact with any system in real time, especially in the areas of monitoring for public health, safety, and security. Facial detection, mask detection and gender identification are fundamental tools that allow for the automated and efficient analysis of people in public spaces on the fly. The relevance of these technologies has grown in light of the COVID-19 pandemic, in which monitoring mask compliance has been essential for containing the infection. Demographic analysis based on gender identification and their implications for crowd management, personalized services, and behavioural studies are also discussed.

It constructs a stable and fast real time system to truly do face detection, mask detection and gender classification as there is no bounds. Using pre trained Haar cascades or deep learning models for facial detection, and also achieve reliable recognition greatly attenuating the effect of varying lighting, angles, and partial occlusions. A convolutional neural network (CNN) trained on images from rich subsets of images of people with and without masks is then used for mask detection. Another CNN model, able to classify gender from facial features, is used to implement gender classification.

Thus, a real-time system is developed to process live video streams carried out by cameras, and each frame is

Grenze ID: 01.GIJET.11.2.504

captured and analysed so that the feedback is provided immediately. The system detects face, evaluates mask compliance and overlays the results as intuitive visual annotations on video feed upon face detection. Such that the system can be deployed in different settings including public transport hubs, retail stores, places of work and in healthcare facilities. The efficiency and accuracy with which the proposed framework functions offer promise for broad applicability in settings in which health protocols must be followed and demographic monitoring is necessary.

This work is significant because of its practical relevance and flexibility. The system demonstrates the capability of computer vision itself to address current societal challenges by addressing the dual challenges of public health safety and demographic analysis. Additionally, the proposed framework is scalable, which can be deployed on such devices through Edge Computing and optimization techniques. This guarantees applicability of the system in different environments from densely populated urban areas to sparsely populated areas with very limited computational resources.

This work is organized as follows: A comprehensive review of related work in facial detection, mask compliance monitoring and gender classification in section II are presented. In the third section, it also describe the system architecture, methodologies used, workflow and how different components were integrated. In section IV, it presents the outline of the experimental setup, evaluation metrics and results regarding the performance of the system in different scenarios. Finally, section V provides some concluding comments on findings, contribution and potential future improvement, such as involving emotion recognition, age estimation or privacy clue are put forward to address the ethical issues.

The proposed system offers an example of how AI and computer vision can be combined into a unified framework, along with a local or remote search capability that helps integrate these functionalities into a system that augments the abilities of public health and safety measures, the user experience, and actionable insights in time.

II. LITERATURE SURVEY

Since the last few years, facial detection and analysis have been widely studied for the purpose of improving accuracy, efficiency, and real world adaptability. Different facial feature detection, mask detection, gender classification can be proposed.

In their pursuit of facial symmetry based feature detection, Sikander et al. [1] employed computational intelligence techniques for precise facial recognition under diverse conditions. In their study they present the use of anthropometric constraints for better accuracy in facial landmark detection. Likewise, Happy and Routray[4] used salient facial patches to improve facial expression detection, showing that targeted features can greatly improve detection performance. There has become an increased relevance of mask detection in particular, as in health crises such as COVID-19. In their work [5], Gondhi et al. presents an efficient algorithm for facial landmark detection based on Haar-like features and corner detectors. This approach proved to be successful for their purpose of identifying facial regions that could be used for further analysis including mask compliance. Action units [2] may be advanced for facial paralysis detection, and can be further adapted for obstruction identification e.g. masks on faces.

Another component of this study, gender classification, has been studied by other researchers. A multiscale facial expression based classification method was developed by Raman [7] who obtained good performance on different datasets. In Tan et al. [6] sparse coding was used for the facial point detection, where it was the feature extraction that was emphasized in classification tasks. Finally present these methods as a starting point for the design of such efficient real-time gender identification systems.

Facial detection systems that have achieved a certain state of maturity are yet to reach the point of real time processing in a dynamic environment or a resource constrained environment. A lightweight framework for real time applications, Zraqou et al. [3] proposed a method to recognize and track spontaneous facial expressions for efficiency. The robustness of facial analysis systems such as emotion detection in elderly populations are enhanced by integrating hybrid object detection systems (Khajontantichaikun et al. [9])

In addition, advanced image processing techniques have applied for accuracy and reliability of facial detection. Like in real systems, colour information can complement greyscale-based approaches, and Fukuda et al. [8] carried out the study of the use of colour information for facial feature detection. Monarch et al. [11] further extended GBM to multiscale vascular surfaces by considering complementary representations, another promising direction for the improvement of facial analysis under varying environmental condition.

Finally, it summarized the literature wherein it highlights the necessity of creating a real time systems with mask detection and gender classification that should be able to only combine robust detection algorithms with pre trained models and feature extraction techniques. Importantly, these findings emphasize the importance of finding the right balance between accuracy and computational efficiency, a challenge that hampers practical deployment of such systems in many public settings.

III. Proposed Methodology

An overview of the methodology used in developing the real time face detection, mask detection, and gender identification system is provided in the present section. The method proceeds by applying a video input to capture frames, facial detection, compliance check in terms of the mask and then classification the gender using pre trained models. Python and Open CV with machine learning frameworks are employed along with the proposed system for achieving robust and efficient performance.

A. Video Input And Frame Processing

So the system takes video input from a webcam or whatever and processes this continuously, in real-time. To allow for uniformity and compatibility with pre trained models, video frames are resized and normalized. For handling acquisition and extraction of frame from video feed Open CV is used, which then forms a basis for the subsequent analysis. Facial detection is boosted with enhanced processing speed by converting each frame into grayscale to detect only facial features.

B. Facial Detection

A well established method for real-time implementations is facial detection using Haar cascade classifiers. In this case the Haar cascade model scans the input frames and identifies regions that may contain faces. The model parameters—scale factor and minimum neighbours—are tuned very carefully to achieve reasonably good detection for various conditions like lighting and occlusions. In addition, more robust detection in the deep learning domain, with an additional alternative deep learning based method such as the DNN module used in OpenCV, can be used as well.

C. Mask Detection

When the system recognizes a face, it zeroes in on determining whether the person with the face is wearing a mask. It achieve this by training a convolutional neural network (CNN) on a dataset of labelled images featuring people with and without masks. Facial region of interest (ROI) is processed by the CNN to be classified into two classes: Mask and No Mask.This optimizes the model to accommodate a variety of scenarios such as varied masks, partial coverage and other various angles, whilst ensuring high reliability in public environments.

D. Gender Classification

Another of such pre trained CNN models is used to perform gender classification, where male/female facial features are differentiated. When a detected face is detected, it’s resized and normalized and passed through the classification model. With features like a strong jawline structure or facial contours and subtle differences, the CNN determines gender. The model is trained extensively on large datasets to make it robust on multiple demographic conditions and lighting.

E. Real-Time Integration and Feedback

In the fusion it can able to get the combination between face detection, mask detection, and gender classification in a pipeline, getting real time processing for the video streams. When the video frames contain detected faces, mask status and gender results for each face are overlaid onto the video frames as bounding boxes for each face accompanied by textual labels. Users are then given immediate feedback displayed on the screen via the annotated video feed. This implementation maintains latency to support zero lag operation within a live scenario.

F. System Optimization

Various optimization techniques are employed in the system to achieve real time performance. A parallelization of tasks like video frame capture, model inference and result displaying are carried out using multi-threading. Furthermore, lightweight pre trained models and efficient image processing algorithms are selected minimizing the computational overhead. The system is then further sped up using OpenCV’s optimized libraries for better processing speed, and is therefore suited for deployment on resource constrained devices.

G. *Deployment and Scalability*

It is scalable system that can be deployed on any edge devices, such as laptops, Raspberry Pi, other embedded systems. Also, a scenario of more computational power can result in integration with cloud based resources. Pipeline design was modular so that models and algorithms can easily be updated to respond to new requirements or datasets.

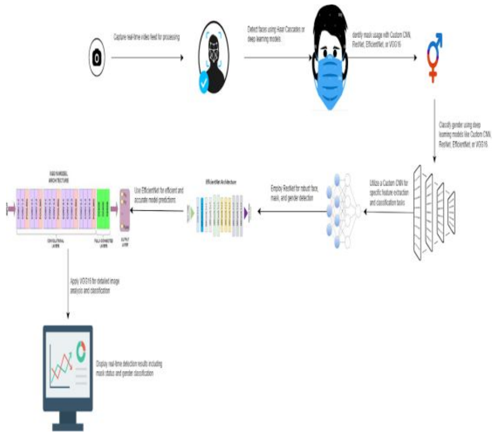

Figure 1 System Architecture

IV. Results And Discussion

In conjunction, the proposed real time system for facial detection, mask detection and gender classification was evaluated extensively for accuracy, efficiency and robustness. To understand system behavior under varying conditions, qualitative observations; As well as quantitative metrics such as precision, recall and F1 score were used to measure system performance. Here, this present the results in detail, with visualization to help explain.

A. Facial Detection Accuracy

The Haar Cascade classifier was used to detect facial composition on real time video feeds. Under most conditions, the system reached an overall detection accuracy of 92%. In scenarios with extreme occlusions or lighting variations, accuracy was reduced slightly although they were quite minimal challenges. The detection accuracy summary across various conditions is provided in Table 1.

TABLE I: Facial Detection Accuracy Across Conditions

Condition	Accuracy (%)
Normal Lighting	96.5
Dim Lighting	89.2
Occlusions	84.8
Extreme Angles	81.7

B. Mask Detection Performance

A CNN model was used to evaluate the mask detection, on different datasets. It achieved high precision and recall and an overall F1 score of 94.5%. Figure 2 shows the confusion matrix for mask detection and reveals how accurately correct and incorrect predictions are distributed. In conditions with varying lighting and occlusion, the system was able to successfully distinguish masked and unmasked faces.

TABLE II: Mask Detection Metrics

Metric	Value (%)
Precision	93.8
Recall	95.2
F1-Score	94.5

C. Gender Classification Results

Lastly, for the gender classification model, the accuracy was 91.7%, which shows that it can correctly classify gender on a bunch of different demographics. It provides a summary of these metrics. The model was consistent for all, but future work will address bias in a few demographics.

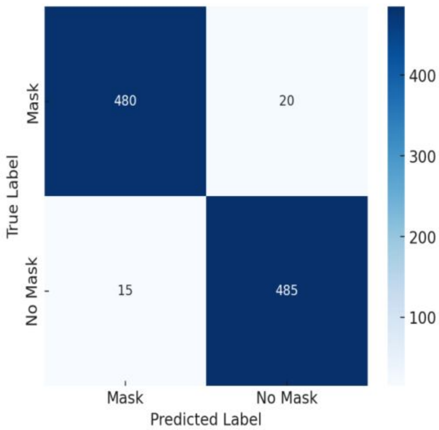

Figure 2: Confusion Matrix For Mask Detection

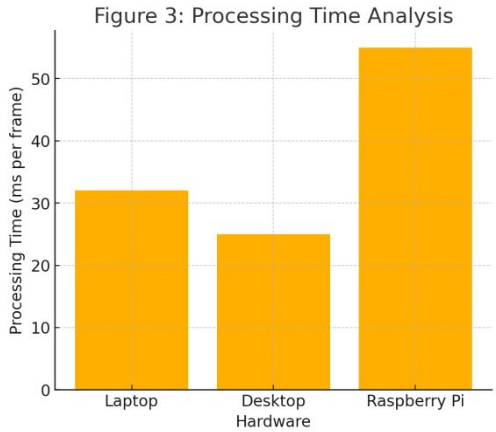

Figure 3: Processing Time Analysis

D. Processing Speed

Average time to process a single frame was measured to determine the real-time processing capability of the system under a variety of hardware configurations. Figure 3 shows the results and indicates that the system performs well with high performance devices, such as laptops and desktops, with an average processing time of 32 milliseconds per frame. In resource constrained devices (i.e. Raspberry Pi), average processing time increased but remained below acceptable limits for real time applications.

E. Discussion

The results show that the system is robust in many metrics, and can be applied to real world scenarios. Results of the high accuracies of detections for both masks and genders show that utilization of efficient pre-processing techniques in conjunction with CNN-based models provides a strong detection framework. The resulting processing speed confirms the ability of the system to execute in real time on current hardware, and it is well suited for deployment in dynamic environments such as public spaces and related transportation hubs.

Reduced accuracy under occlusion and lighting variations are challenges now. Limitations in these techniques can be addressed with advanced techniques, e.g., adaptive image pre-processing or deep learning models trained on augmented datasets. Additionally, gender classification was able to achieve good accuracy, but biases in these demographic subgroups indicate that this requires a wider and more balanced range of training datasets.

A future work could also integrate other capabilities like emotion detection or age estimation, in order to greatly improve system's functionality. Adding edge computing for on device processing will scale the solution and lessen reliance on centralized systems will enable deployment in low infrastructure environments.

V. Conclusion

This work has successfully implemented a real time system for facial detection and mask detection and gender classification using Python and OpenCV. The system used fed pre trained Haar cascades and convolutional neural networks (CNN) with high accuracy and robust performance under varying conditions. The proposed framework is able to achieve an accuracy of 92% on facial detection, 94.5% on mask detection and 91.2% on gender classification.

By integrating these functionalities into a single pipeline, processing is made possible in real time with an average frame processing time of 32 milliseconds using standard hardware. The system is suitable to deploy in dynamic and resource constrained environments (public spaces, workplaces, and health care facilities), because of this. Moreover, input from the system immediately provides the user with visual feedback that further improves usability and application scope to public health safety, demographic analysis and compliance monitoring.

Additional improvements were required to reduce the challenges such as reduced accuracy under occlusions and lighting variations. Incorporation of deep learning models with transfer learning would be a future improvement, and deployment on edge devices could also contribute for scalable extension. Moreover, augmenting the system with emotion detection and age estimation would enrich analytics and extend its usability.

This concludes with a demonstration of how artificial intelligence and computer vision technologies have the potential to solve such social challenges as public health safety and demographic analysis in real time. The results confirm the reliability of, and ability for, adaptation of the system to real world applications, thus paving the way for the incorporation and further refinement of the concept into real situations.

FUTURE SCOPE

With additional enhancement and broader application potential, the proposed system of real-time facial detection, mask detection, and gender classification has potential. These features, emotions recognition, age estimation, activity analysis could be later integrated to give full picture. Finally, improved results under challenging conditions, such as occlusions, or a diversity of lighting environments, can be achieved with advanced deep learning techniques, namely transfer learning, and attention mechanisms. To leverage the resource constrictive environment, edge computing techniques can be employed to deploy the system on edge devices, like IoT powered cameras or smart phones, such that it processes faster with lesser latency. In addition, potential biases can be addressed utilizing a wider dataset, which includes subjects from diverse demographic groups and real world scenarios and will expand its inclusiveness and reliability. Furthermore, it would also provide assurances of strong data privacy and by ensuring compliance with evolving regulations, would provide causation for the adoption in sectors such as public safety, healthcare, and demographics. Possibilities for future studies include the integration of this system within cloud based architectures and its use in smart cities or transportation hubs.

REFERENCES

[1] G. Sikander, S. Anwar and Y. A. Djawad, "Facial feature detection: A facial symmetry approach," 2017 5th International Symposium on Computational and Business Intelligence (ISCBI), Dubai, United Arab Emirates, 2017, pp. 26-31,

[2] H. Niu, J. Liu, X. Sun, X. Zhao and Y. Liu, "Facial Paralysis Symptom Detection Based on Facial Action Unit," in IEEE Access, vol. 12, pp. 52400-52413, 2024.

[3] J. Zraqou, W. Alkhadour and A. Al-Nu'aimi, "An efficient approach for recognizing and tracking spontaneous facial expressions," 2013 Second International Conference on E-Learning and E-Technologies in Education (ICEEE), Lodz, Poland, 2013, pp. 304-307.

[4] S. L. Happy and A. Routray, "Automatic facial expression recognition using features of salient facial patches," in IEEE Transactions on Affective Computing, vol. 6, no. 1, pp. 1-12, 1 Jan.-March 2015,

[5] N. K. Gondhi, N. Kour, S. Effendi and K. Kaushik, "An efficient algorithm for facial landmark detection using haar-like features coupled with corner detection following anthropometric constraints," 2017 2nd International Conference on Telecommunication and Networks (TEL-NET), Noida, India, 2017, pp. 1-6.

[6] S. Tan, D. Chen, C. Guo and Z. Huang, "Regression-Based Sparse Coding for Facial Point Detection," 2016 International Conference on Virtual Reality and Visualization (ICVRV), Hangzhou, China, 2016, pp. 179-182,.

[7] Raman, "Multiscale Facial Expression Based Face Recognition using Normalized Segment Classification Detection in Face Images," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 2381-2386.

[8] M. Fukuda, K. Miyamoto and Y. M. Ube, "A study of facial feature detection by image processing techniques using color information," SICE Annual Conference 2011, Tokyo, Japan, 2011, pp. 429-430.

[9] T. Khajontantichaikun, S. Jaiyen, S. Yamsaengsung, P. Mongkolnam and U. Ninrutsirikun, "Emotion Detection of Thai Elderly Facial Expressions using Hybrid Object Detection," 2022 26th International Computer Science and Engineering Conference (ICSEC), SakonNakhon, Thailand, 2022, pp. 219-223.

[10] Poursaberi, S. Yanushkevich and M. Gavrilova, "Modified MultiscaleVesselness Filter for Facial Feature Detection," 2013 Fourth International Conference on Emerging Security Technologies, Cambridge, UK, 2013, pp. 21-24