Environmental Risk Handling on Face Landmark Detection Using Computer Vision

K. Gayathri

B.E., M.E., PhD Scholar, Information Technology, Anna University, Velammal College of Engineering and Technology, Madurai, India
Email: *jerrygayu[at]gmail.com*

Abstract: ***Now a days we are facing environmental risks on face landmark detection and it is a more important task thus it can be rectified using simple hardware such as computer vision for widespread applications such as face recognition, expression analysis, and human computer interaction. Visualization to demonstrate face landmark detection on huge environmental conditions remains a goal challenges. This paper presents a lightweight computer vision pipeline that integrates Media Pipe's, Face Mesh with OpenCV preprocessing techniques to enhance robustness against pose variation, lighting inconsistency, and occlusion. The proposed method classify 3 stages pipeline on histogram equalization, gamma correction, Haar cascade classifier for face detection, and edge-based occlusion to improve landmark detection as real world unconstrained environments. Experimental result shows correct effective accuracy of face landmark as face mesh 90% face landmark detection over real-world to avoid face in security against environmental risk and make world more secure using over face Technology.***

Keywords: Face Landmark Detection, MediaPipe, Face-mesh OpenCV, Pose Variation, Lighting Correction, Occlusion Detection, Real-Time Vision

1. Introduction

Face alignment is a process of identifying human face features and it aims to classify a set of predefined human facial landmarks such as the corners of the eye, eyebrows, and nose, mouth edges of face for high-level tasks such as facial recognition, facial point matching, facial animation, and 2d& 3d facial modeling. Face recognition emerged to identify human face by using several computation and has witnessed marked improvement, it has attracted a lot of interest in the past couple of years. Understanding this tendency is possible for at least two reasons: The accessibility of the pertinent technology (such as cell phones, digital cameras, GPU), in addition to the wide range of commercial and legal needs.

Machine learning on face recognition systems have improved massively in the current century, but their performance is still limited by the requirements of real-world settings and applications. Major risks such as multiple obstacles in detecting face photos taken in an unrestricted environment (such as changes in light exposure, sitting posture, or facial expression, including partial occlusion, disguises, or lens movement). Face landmark detection plays a crucial role in applications like human face emotions, face recognition, and face virtual avatars, and for verification purposes, face biometric systems are used.

Recently environmental challenges have played a role as varying lighting conditions, occlusions, and head pose angles can severely impact detection and give 100% accuracy. Deep learning-based solutions often require high computational power by calculating huge datasets while in limiting their user's flexibility in real-time environment by using mobile applications. Face mesh is one of the new advanced method which identify whole face with 468 3D landmarks over eyes, eyebrows, nose, mouth, lips, jawlines and more thus it can be analyzed using real time environment using media pipe by Google and these face mesh calculates 2 dimension and 3dimensional faces over the region and thus process can be done using media pipe Google.

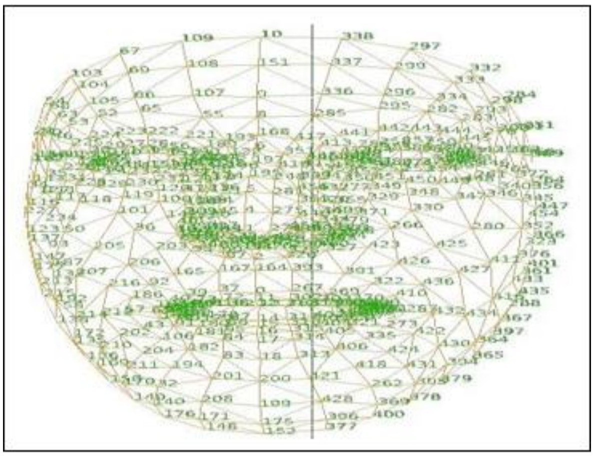

Figure 1: Face mesh

2. Literature Survey

a) *Deep Learning Approaches*

Zhang et al. [1] approach based on a Deep learning technique using 2 paired (CS) cascaded subnetworks with CNN units 1st network estimate heatmap-based encodings of the landmarks' locations, 2nd network receive as inputs the outputs of heatmap estimation units, and classify to regression. Chang et al. introduced FacePoseNet [6] shows better landmark detection accuracy on the 300W benchmark has better face recognition accuracy. Here FPN is used for superior 2D and 3D face alignment in small fraction of the computational cost of comparably accurate landmark detectors. Pixel-in-Pixel Net introduced by Jin et al. [7], proposed process detect face head based on heatmap regression over pixel in pixel network checks score and offset predictions on low-resolution on face feature maps and repeating on sampling layers sacrificing model accuracy with huge dataset. Haoqi and Koichi [8]. The effective neighbor regression module is proposed to enforce local constraints by fusing predictions from neighboring landmarks develops to enhances the robustness on dark region preprocessing classified using heatmap of face develop the face region by cropping the features face and extracting face region by various algorithm.

b) Multi-View and Dense Landmark Models

Deng et al. [9], results on videos captured and dataset of animals consisting the image frames on their corresponding 2D and 3D labels for finding edges of images normally in rigid geometry scale detection. Retina Face by Deng et al. [10] proposed system is done on uncontrolled face detection which accurate the efficient face localization as wild and remains an open challenge in robust single-stage face detector that performs Google net

c) Lightweight and Real-Time Systems

Google's MediaPipe framework [2] MediaPipe is Google based open source performs framework over rapid combination existing new perception components of prototypes for advance crossplatform applications. Developer can easily built MediaPipe to manage resources on both CPU and GPU for less latency performance and to handle real time-series data such as audio, video frames and to measure performance. [3], it also show features extraction of various images identification enabling over various technology for program developer to focus on the algorithm based or model-based development using MediaPipe as an environment for iteratively improvement and as open source platform and get coding easier in GITHUB.

d) Classical and Hybrid Models

Viola and Jones [4] Developed hybrid algorithms for multistage face alignment that has potential face detection particular at each model optimization and machine learning to integrate methodologies in huge networks with advance deep neural networks. This paper presents an extensive systematic and biometric literature review on hybrid face identification methods. [3] Kahou et al.'s face rotation binding for finding faces involves optimization on algorithm- based techniques for clustering and classification of image identify the potential models as confident accuracy methodologies using huge compared dataset with CNN regression- based image classification for combined digital computation with networks and IOT based identification

3. Existing System

Past approaches primarily rely on convolutional neural networks (CNNs), Google net, Resnet and large pose occlusion face datasets they are not determined properly. Method is more accurate, they are computationally intensive. Media Pipe develop face images identification on real live faces with faster, low resolution, resource-friendly alternative using optimized models. Previous works have explored lighting normalization and pose compensation independently, but here few challenges are found clearly all three major risks are preprocessed over classification of face pose, lighting, and occlusion are not identified.

Challenges	Technique Used	Purpose
Lighting Correction	Histogram Equalization, Gamma Correction	Normalize illumination
Pose Handling	Haar Cascade Face Detection	Locate and align face region
Occlusion Awareness	Canny Edge Detection	Visual cue for obstruction
Lightweight Processing	MediaPipe, OpenCV	Fast execution, mobile-capable

4. Proposed System

The proposed system mainly process 3 pipeline frames perfectly while on challenges here we used **lightweight, modular face landmark detection pipeline to reduce accuracy and faster processing on face landmark detection as fraud detection** designed to work reliably in real-world, unconstrained environments. It combines traditional computer vision preprocessing with Media Pipe's for efficient FaceMesh to address key environmental risks like lighting variation, pose changes, and occlusion.

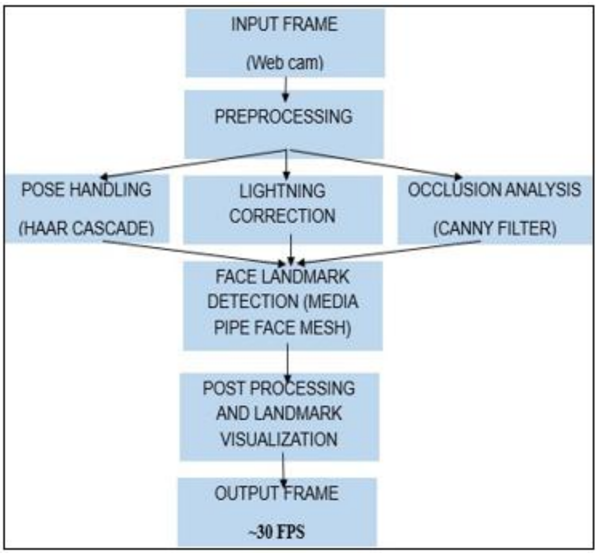

Figure 2: Proposed System

5. Methodology

1) System

Here system or laptop captures video frames from a webcam. Each frame pipeline is stages to preprocessing and handle pose variation, lighting inconsistency, and occlusion before passing to the landmark detection stage as after preprocessing. The processed frame is then analyzed using a facial landmark detection model to extract meaningful features.

2) Preprocessing

a) Pose variation
b) Lighting inconsistency
c) Occlusion analysis

a) **Pose Variation Handling**: Haar cascade classifiers (HCC) finds regions and classified properly to localize processing and provide a bounding box for landmark detector to improve efficiency and providing spatial constraints for drawing on face landmark overlay. Haar

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

cascade classifiers (HCC) is a main machine learning process that detects object and used to identify objects in various images or video and widely used for **face detection**. It works by applying a simple rectangular features (called **Haar-like features**) to a grayscale images which are evaluated quickly using an **integral images.** Haar cascade is a new and faster method of pre-processing images over 3 stage pipeline and 1st it checks haar algorithm to detecting faces region it covers rectangular shape over face features as to identify edge features, line features and four rectangle features that convert as rectangular features of face determines which are similar to the kernel processing detection on face regions.

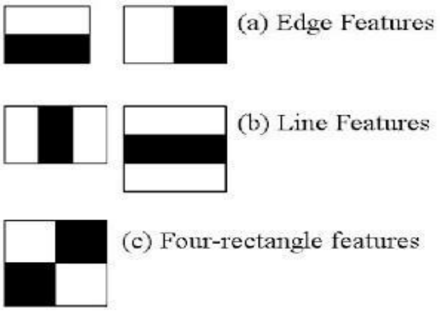

Figure 3: Haar cascade classifiers

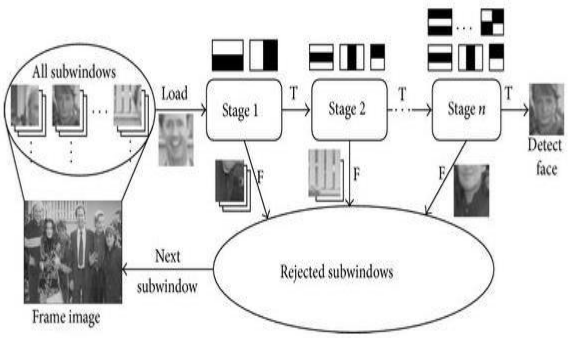

Figure 4: Working of Face Detection using

HAAR Cascade

- **Step 1:** The images has been sent to the classifier that divided into small parts in sub windows
- **Step 2:** N number can be detected using cascading classifier learns a combination of different types of features from images like line, edge, circle, and square that are passes on classification of image . While feature extraction is done each sub-part is assigned a confidence value.
- **Step 3:** Images or sub-images has highest confidence and detected as face and sent to
- the accumulator while the rest are rejected finally cascade fetches nearest frame of image if remaining and starts the processing again.

b) **Lighting Correction:** Histogram equalization is applied in grayscale space to normalize lighting over face region.

L: total number of possible intensity levels (usually 256 for 8-bit images)

r_k: the original intensity level (from 0 to $L-1$)

n_k: number of pixels with intensity r_k

MN: total number of pixels in the image (M rows × N columns)

1. **Probability of intensity level:**

$$p(r_k) = \frac{n_k}{MN}$$

2. **Cumulative Distribution Function (CDF):**

$$cdf(r_k) = \sum_{j=0}^{k} p(r_j)$$

3. **Transformation Function:**

$$s_k = \text{round}\left[(L-1) \cdot cdf(r_k)\right]$$

- s_k is the new intensity value
- L- 1 is the maximum possible intensity value (255 for 8-bit)

4. Complete Mapping

Each pixel T_k is mapped to s_k, producing a new image with enhanced contrast

Summary Equation

$$s_k = (L-1)\sum_{j=0}^{k} \frac{n_j}{MN}$$

Gamma correction on visual consistency across frames while illumination.

Occlusion Analysis: Edge detection using canny filters visualizes possible occlusions. Although not currently used to mask the landmark model, it aids in future occlusion-aware enhancement.

1) **Apply Gaussian Filter (Noise Reduction)**

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

This smooths the image to reduce the impact of noise before edge detection

2) **Compute Image Gradients (Edge Strength & Direction)**

Using Sobel Operators

$$G_x = \frac{\partial I}{\partial x}, \quad G_y = \frac{\partial I}{\partial y}$$

Then Calculate

- Gradient Magnitude

$$M(x,y) = \sqrt{G_x^2 + G_y^2}$$

- Gradient Direction

$$\theta(x,y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

3) **Non- Maximum Suppression**

Thins the edges by keeping only local maxima of *M (x, y)* along the direction Ø *(x, y)*

For a normalized input pixel value I_{in} **ε** [0, 1], the gamma-corrected output I_{out} is:

$$I_{out} = I_{in}^{\gamma}$$

Or, more commonly in image processing:

$$I_{out} = 255 \times \left(\frac{I_{in}}{255}\right)^{\gamma}$$

Where

- $\gamma < 1$: brightens the image (useful for underexposed images)
- $\gamma > 1$: darkens the image (for overexposed images)
- $\gamma = 1$: no change

4) Double Thresholding

Classify image pixels

- Strong edges: $M(x, y) > T_{\text{high}}$
- Weak edges: $T_{\text{low}} < M(x, y) \leq T_{\text{high}}$
- Non-edges: $M(x, y) \leq T_{\text{low}}$

5) Hysteresis (Edge Tracking)

Connect Weak edges to strong ones to preserve true edges and discard isolated noise

c. Landmark Detection:

MediaPipeFaceMEsh is employed with a detection confidence threshold of 0.5. Detected 468 landmarks and drawn using predefined facial contour connections for better integrations.

D. Implementation: The entire pipeline is implemented in Python using OpenCV and MediaPipe. Webcam input is processed in real-time (~30 FPS). Each stage pipeline is modular and adjustable for different device capabilities usage to be tuned independently for performance.

E. Discussion: The system successfully detects landmarks under varying lighting and head orientation. Edge-based occlusion reveals partial obstruction areas. Then edges are Compared to baseline detection without preprocessing, the system improves consistency and detection under sever conditions are supported by both qualitative observations and quantitative analysis.

6. Algorithm

Input: Real-time video stream from the webcam of the computer

Output: visualize facial landmarks with robustness like pose, lighting, and occlusion identification

Step1: open python webcam opencv media pipe face mesh with confidence thresholds **Step 2**: Each frame pipeline that capture from the webcam:

- Convert image to grayscale
- Apply histogram equalization to normalize Lighting
- Apply gamma correction to adjust lighting
- Using Haar cascade algorithm to detect the face boundary box
- Apply canny edge detection for occlusion visualization
- Feed the original frame to media pipe face mesh
- Extract 468 facial landmarks and contours using opencv
- Draw landmarks and contours using opencv
- Display the final frame with overlays
- Release the webcam and close all windows upon exit.

7. Result

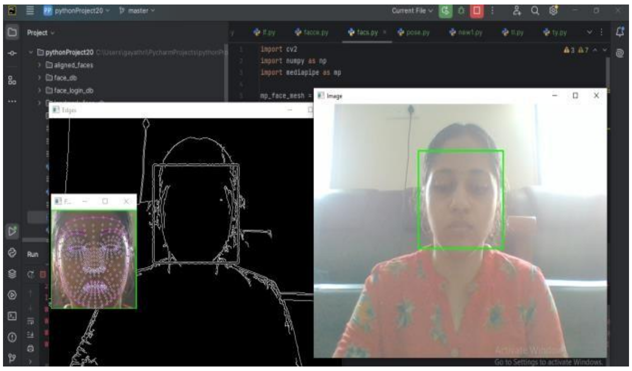

Figure 5: Real time face landmark detection using computer vision identifies image of face with boundary mark, edges of face with boundary and face mesh of analyzed face

8. Experiment

The proposed system was evaluated under various environmental conditions, including variations in face lighting, face of head orientation, and partial facial occlusion. Hardware configuration as Intel i5 processor and 8 GB of RAM for real-time video input captured via a webcam can be used for lightweight capturing

The experimental outcomes are summarized below:

1) Computation of face analysis **90% accuracy** on facial landmark detection under moderate lighting and occlusion risks
2) While integrating preprocessing techniques on MediaPipe of FaceMesh on the system results a high **detection performance of 95%**.
   - Finally the pipeline of frame maintained a consistent **processing speed of approximately 30 frames per second (FPS), confirming its suitability for realtime applications.**
   - Compared to the baseline performance using MediaPipe FaceMesh alone and the proposed approach demonstrated **(12– 15%) improvement in detection stability** handle for challenging environment risks.

Table 1: Comparison table for Face Landmark Detection Performance

Performance Metric	MediaPipe	Proposed Pipeline	Improvement
Accuracy (Moderate Lighting & Occlusion)	78%	90%	+12%
Overall Detection Performance	88%	95%	+7%
Detection Stability in (Challenging Conditions)	low	high	+13.5%(avg)
Real-Time Speed	~30 FPS	~30 FPS	No change

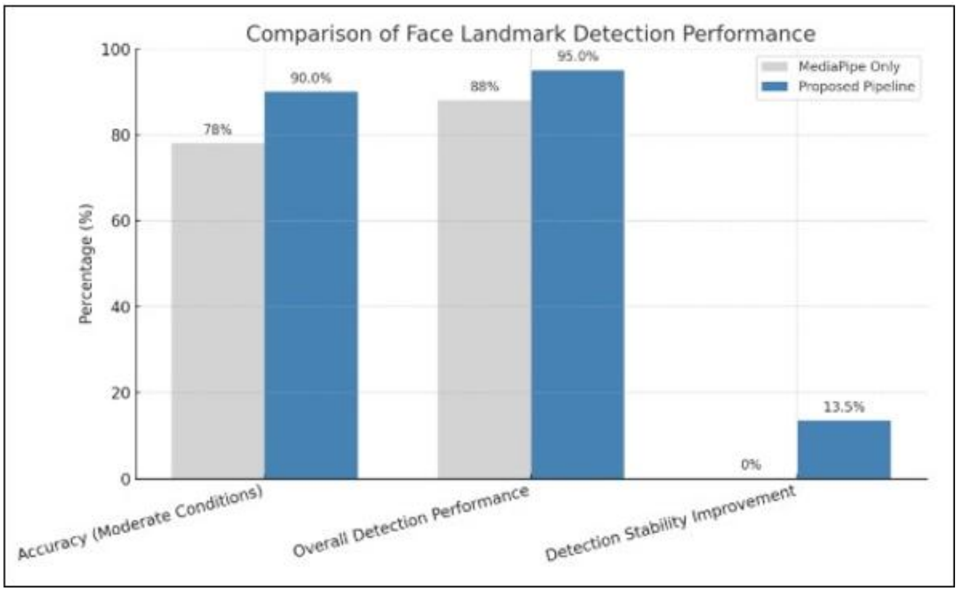

Figure 6: Comparison chart

9. Conclusion

In this paper we results effective approach by developing consistent face landmark accuracy on real-time risk over environment for securing purpose . The proposed system detects lightweight, real-time face landmark detection system over the face region that relies on challenges to integrate difficulties in pose normalization, lighting correction, and occlusion visualization, this system enhances these processes by using media pipe, robustness in varied pipeline scenarios it shows extract 468 facial landmark of face mesh by using OpenCV Media pipes. Future development explore automated occlusion-aware over face filtering and thus it can be viewed over mobile application oriented platform for various real-world applications and to make more awareness on face detection against robust AI classification on deep learning to identify malware detection over cyber security.

References

[1] F. Zhang, X. Liu, and Z. Zhang, “Detecting Facial Landmarks via Cascaded Deep Convolutional Networks,” IEEE Transactions on Image Processing, 2016.

[2] MediaPipe by Google: https://mediapipe.dev/ [3] G. Bradski, “The OpenCV Library,” Dr. Dobb’s Journal of Software Tools, 2000.

[3] P. Viola and M. Jones, “Robust Real-Time Face [6] F. J. Chang, A. Tuan Tran, T. Hassner, I. Masi, R. Nevatia and G. Medioni, "Faceposenet: Making a case for landmark-free face alignment", Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 1599-1608, 2019.

[4] H. Jin, S. Liao and L. Shao, "Pixel-in-pixel net: Towards efficient facial landmark detection in the wild", International Journal of Computer Vision, vol. 129, pp. 3174-3194, 2021

[5] G. Haoqi and O. Koichi, "Improvements over Coordinate Regression Approach for Large-Scale Face Alignment", IIEEJ Transactions on Image Electronics and Visual Computing, vol. 10, pp. 127135, 2022.

[6] J. Deng, G. Trigeorgis, Y. Zhou and S Zafeiriou, "Joint multi-view face alignment in the wild", IEEE Transactions on Image Processing, vol. 28, pp. 3636-3648, 2019.

[7] J. Deng, J. Guo, E. Ververas, I. Kotsia and S. Zafeiriou, "Retinaface: Single-shot multi- level face localization in the wild", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5203-5212, 2020 Detection,” International Journal of Computer Vision, 2001.

[8] S. E. Kahou et al., “Combining modality specific deep neural networks for emotion recognition in video,” ACM ICMI 2013.