International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

FACIAL LANDMARK DETECTION WITH MEDIAPIPE & CREATING ANIMATED SNAPCHAT FILTERS

M.Meghana[1], M.Vasavi[2], Dr. D. Shravani[3]

Department of Computer Science and Engineering, Stanley College of Engineering and Technology for Women, Telangana, India

Abstract. MediaPipe Face Mesh is a face geometry solution that estimates 468 3D face landmarks in real-time even on mobile devices. It employs machine learning (ML) to infer the 3D surface geometry, requiring only a single camera input without the need for a dedicated depth sensor. Utilizing lightweight model architectures together with GPU acceleration throughout the pipeline, the solution delivers real-time performance critical for live experiences. There are two models and they are face detection and face landmark model.Face landmark screen coordinates are converted into the Metric 3D space coordinates; face pose transformation matrix is estimated as a rigid linear mapping from the canonical face metric landmark set into the runtime face metric landmark set in a way that minimizes a difference between the two; a face mesh is created using the runtime face metric landmarks as the vertex positions

(XYZ), while both the vertex texture coordinates (UV) and the triangular topology are inherited from the canonical face model.Snapchat Filter Controlled by Facial Expressions: Currently, our face expression recognizer can check whether the eyes and mouth are open or not so to get the most out of it, we can overlay scalable eyes images on top of the eyes of the user when his eyes are open and a video of fire coming out of the mouth of the user when the mouth is open.

Keywords: Computer vision, neural networks, mobile applications face recognition, face Land marking, Registration, Image, Vision, Machine learning, Deep learning, Review, Survey.

1. Introduction

1.1 About Paper

OpenCV is the cross-platform open-source library for computer vision, machine learning, and image processing using which we can develop real-time computer vision applications. It is mainly used for image or video processing and also analysis including object detection, face detection, etc.Facial landmarks are used to localize and represent important regions of the face, such as:

- Mouth
- Eyes
- Eyebrows

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

- Nose
- Jawline etc.

Facial landmark detection/estimation is the process of detecting and tracking face key landmarks (that represent important regions of the face e.g, the center of the eye, and the tip of the nose, etc) in images and videos. It allows you to localize the face features and identify the shape and orientation of the face.Snapchat Filter Controlled by Facial Expressions: Currently, our face expression recognizer can check whether the eyes and mouth are open or not so to get the most out of it, we can overlay scalable eyes images on top of the eyes of the user when his eyes are open and a video of fire coming out of the mouth of the user when the mouth is open.

1.2 Objectives of the Paper

This project aims to design a system that takes a video as an input splitting it into frames and obtains images extracting its features and landmarks through the algorithms of Machine learning thus providing a base for a number of possible systems. The main Objectives of this project are:

- To design a system that takes an image as an input and detects the facial landmarks including lip movement tracking and lip gesture recognition system
- To study the application of the system in automatic voice recognition as incorporating audio and visual information together can improve the performance.
- To analyze the scope of the system in face detection that is a key step in any kind of facial recognition system.
- To research the possibility of implementation of the system thus developed in assistive technology for differently abled people such as controlling a system or program through the movement of lips.

1.3 Scope of the Paper

Facial landmark detection algorithms help to automatically identify the locations of the facial key landmark points on a facial image or from a video. The key landmark points normally includes the facial regions like nose tip, eye corner, eye brows and chin tip. Some applications of facial landmark detection are face swap, head pose detection, detecting facial gestures, gaze direction etc. Face detection locate a human face from the image and return a value in terms of the co-ordinates of the bounding rectangle of the detected face. After detecting the face, we have go through points inside the bounding rectangle for key points.

Filters come in a variety of filters. Facial filters are a form of virtual reality. wherein facial recognition creates a digital animation and alters a user's reality on-screen.

2. Literature Survey

2.1 Existing System

The existing system uses Digital camera for facial recognition.

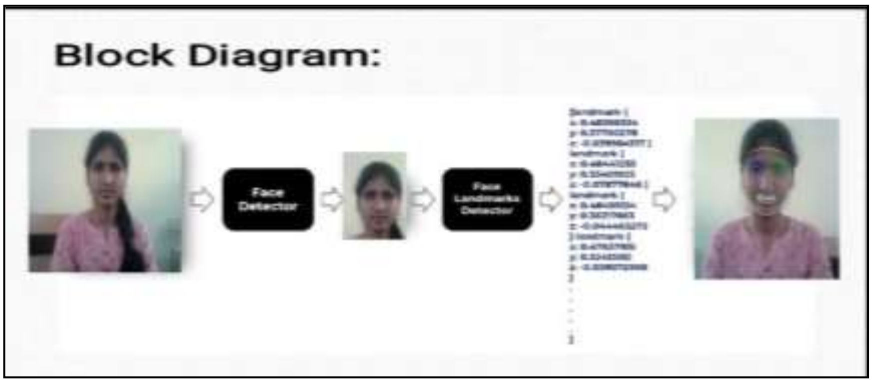
The camera was used as an input device and the object need to present in front of the camera for capturing Facial landmark.

It limits mobility and cannot use camera in dark area as well as cost is very high**.**

The pipeline is implemented as a MediaPipe graph that uses a face landmark subgraph from the face landmark

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

module, and renders using a dedicated face renderer subgraph. The face landmark subgraph internally uses a face_detection_subgraph from the face detection module.



2.2 Proposed System

To precisely estimate the position of facial landmarks including lip coordinates in a computationally efficient way using the algorithms of machine learning.It is expected to achieve an accuracy that is feasible in a technical as well as economical manner

The project will also focus on obtaining the information on how well it suits for different applications.

The purpose of face alignment is to locate semantic facial landmarks automatically and to rectify face images into the same canonical pose (typically, the front view) which is essential for some tasks. For instance, expression recognition , face swapping , head pose estimation, face tracking , face animation , blink detection and 3D face modelling.

3. Proposed Architecture

3.1 BLOCK DIAGRAM

Fig 3.1: Block Diagram of 3d Landmark Coordinates

The machine learning pipeline of the Mediapipe's solution contains two different models that work together:

- A face detector that operates on the full image and locates the faces in the image.
- A face landmarks detector that operates only on those face locations and

predicts the 3D facial landmarks.

So the landmarks detector gets an accurately cropped face ROI which makes it capable of precisely working on scaled, rotated, and translated faces without needing data augmentation techniques.

In addition, the faces can also be located based on the face landmarks identified in the previous frame, so the face detector is only invoked as needed, that is in the very first frame or when the tracker loses track of any of the faces.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

3.2 BLOCK DIAGRAM OF FACE LANDMARKS

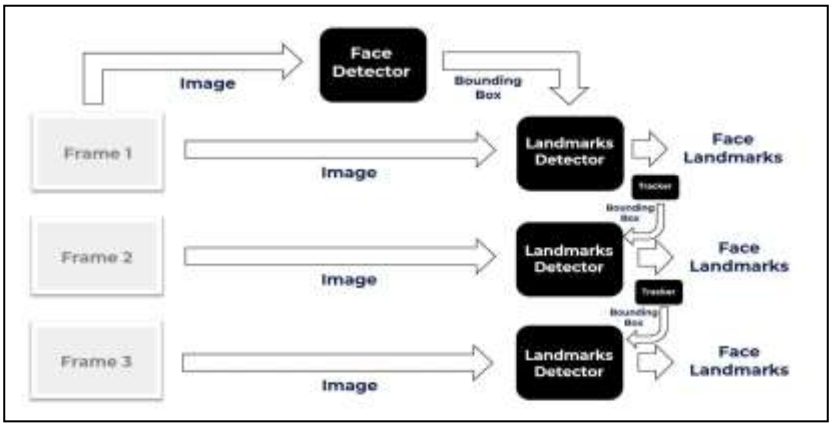


Fig 3.2: Block Diagram of face landmarks

They have utilized transfer learning and used both synthetic rendered and annotated real-world data to get a model capable of predicting 3D landmark coordinates. Another approach could be to train a model to predict a 2D heatmap for each landmark but will increase the computational cost as there are so many points.

3.3 FACIAL LANDMARK DETECTION

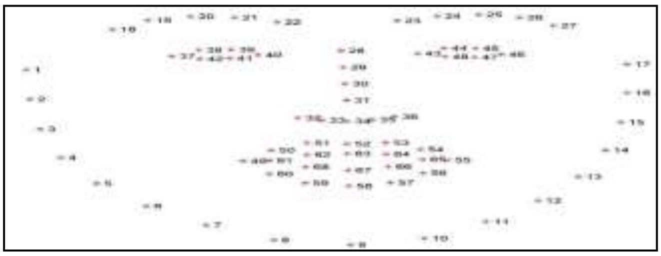


Fig 3.3: Facial Features points

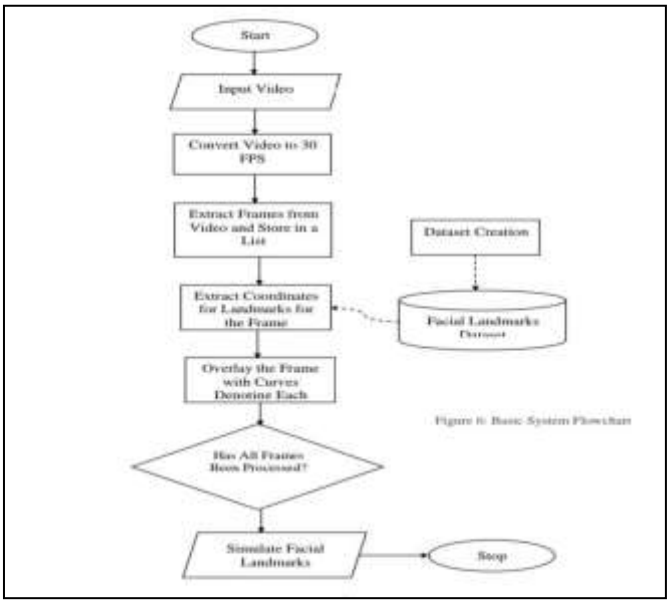
Facial feature points are termed as landmarks or facial fiducial points. These points are principally situated around edges or corners of facial parts, for example, eyebrows, eyes, mouth, nose, and jaw. Existing databases available for facial landmark detection are labelled with varying number of feature points, differing from the 5-point configuration to the maximal 194-point configuration. For the most part, facial landmark detection is a supervised or semi-supervised learning process that trains on a substantial number of well-marked facial pictures. Basically, facial landmarks detection

begins from a face detection process and after that prediction of the facial landmark, points is carried out inside the identified face bounding box. The confined facial landmark points can be used for different face analysis assignments, for example, face recognition, facial animation, facial expression detection, and head pose .

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

3.4 SYSTEM FLOWCHART

Fig 3.6: Basic system flowchart



4.1 Algorithm

The Facial appearance refers to the distinctive pixel intensity patterns around the facial landmarks or in the whole face region, while face shape patterns refer to the patterns of the face shapes as defined by the landmark locations and their spatial relationships.

1. Mediapipe Model :

   Mediapipe is a cross-platform library developed by Google for computer vision tasks. Mediapipe python library uses a holistic model to detect face and hand landmarks. This holistic model produces 468 Face landmarks, 21 Left-Hand landmarks, and 21 Right-Hand landmarks. The individual landmarks can be accessed by specifying the index of the required landmark. Unfortunately, mediapipe library now supports only video stream elements, not functioning on static images.

2. Convolutional Neural Network(CNN):

   CNN is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

   The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

3. Support Vector Machine:

Support Vector Machines are linear classifiers that maximise the margin between the decision hyperplane and the examples in the training set. Osuna et al. first applied this classifier to face detection.

4. Adaboost

   The AdaBoost-based face detection algorithm is an algorithm that is based on integral image, cascade classifier and AdaBoost algorithm. The basic idea is as following: First it uses the integral image to calculate face's Haar-like features quickly.

4.2 Code Implementation

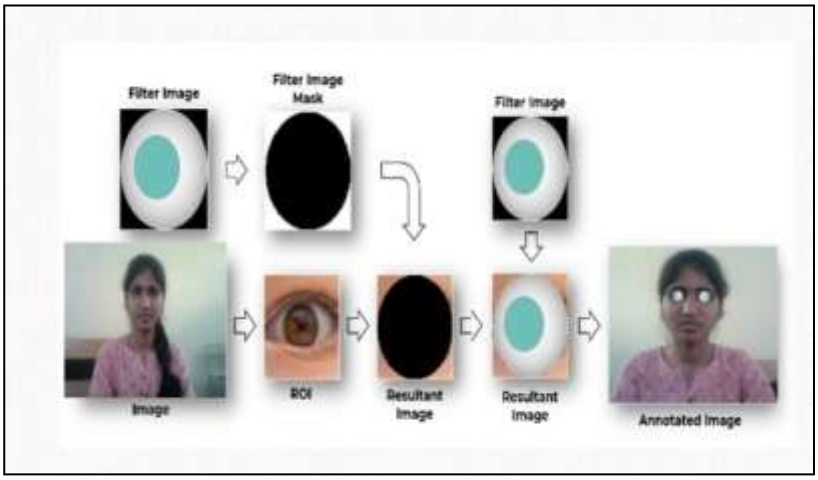
Tensorflow. TensorFlow is an amazing information stream in machine learning library made by the Brain Team of Google and made open source in 2015.It is intended to ease the use and broadly relevant to both numeric and neural system issues just as different spaces. Fundamentally, TensorFlow is a low level tool for doing entangled math and it targets specialists who recognize what they're doing to construct exploratory learning structures, to play around with them and to transform them into running programs.

Anaconda3 5.3.1. Anaconda is a free and open-source appropriation of the Python and R programming for logical figuring like information science, AI applications, large-scale information preparing, prescient investigation, and so forth. Anaconda accompanies in excess of 1,400 packages just as the Conda package and virtual environment director, called Anaconda Navigator, so it takes out the need to figure out how to introduce every library freely. to Anaconda appropriation that enables clients to dispatch applications and oversee conda packages, conditions and channels without utilizing command line directions.

Mediapipe: Mediapipe is a cross-platform library developed by Google for computer vision tasks. Mediapipe python library uses a holistic model to detect face and hand landmarks. This holistic model produces 468 Face landmarks, 21 Left-Hand landmarks, and 21 Right-Hand landmarks. The individual landmarks can be accessed by specifying the index of the required landmark. Unfortunately, mediapipe library now supports only video stream elements, not functioning on static images.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

5.Result



International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

6. Conclusion

From a detailed survey of facial landmark detection algorithms, we make the following conclusions: despite a significant growth of methods' quality, few of them focus on the real-world applicability, meaning that in many cases even when executed on a GPU, algorithms performs slower that real-time (around 30 fps or33 milliseconds).Snapchat filters are fascinating progressions within the scope of augmented reality. The developments that have been applied to the scope of calibration and augmenting of real and fake together into a single application in the palm of one's hand has allowed for interesting deviation of reality to exist. Snapchat filters are fascinating progressions within the scope of augmented reality. The developments that have been applied to the scope of calibration and augmenting of real and fake together into a Single application in the palm of one's hand has allowed for interesting deviation of reality to exist. As Snapchat grows in popularity, the quality and variations of the augmenting of what the lens captures expands constantly.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

7. Future Scope

The quantity of facial appearances and their variability has played an instrumental role in the learning of significant representations and improving the fitting performance. Dense 3D reconstruction instead of relying on only sparse landmarks can be done with the help of optical flow which is able to track dense points on the face in the image space. Facial features components can be characterized as holistic in both shape and texture parameters, these features must be strengthened by fusing of other representations such as capitalizing on both geometry and texture while applying the 3D reconstruction, or adding sparse learning into the 3DMM fitting to achieve more robust and accurate results. It would be useful to examine more robust and precise fusion strategies

8. References

1. **1**.https://towardsdatascience.com/face-landmark-detection-usingpython 1964cb620837 #:~text:=Face%20landmark%20detection%20is%20a,is%20paying%20attention%20or%20not.
2. **2.**https://pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/
3. **3.**https://jivp-eurasipjournals.springeropen.com/articles/10.1186/1687-5281-2013-13
4. **4.**https://medium.com/codex/facial-landmark-detection-algorithms-5b2d2a12adaf
5. **5**.https://arxiv.org › pdf Fast Facial Landmark Detection and Applications: A Survey – arXiv
6. **6.**https://www.researchgate.net › 3464 Fast Facial Landmark Detection and Applications: A Survey
7. **7.**jivp-eurasipjournals.springeropen.com A review of image-based automatic facial landmark identification …
8. **8.**https://medium.com › codex › facial…Facial Landmark Detection Algorithms | by Raji Lini | CodeX
9. Kishor Kumar Reddy C, Rupa C H and Vijaya Babu B, "SLGAS: Supervised Learning using Gain Ratio as Attribute Selection Measure to Nowcast Snow/No-Snow", International Review on Computers and Software, 2015.

10. (http://www.praiseworthyprize.org/jsm/index.php?journal=irecos&page=article&op=view&path%5B%5D=16706)
11. Kishor Kumar Reddy C, Vijaya Babu B, Rupa C H, "SLEAS: Supervised Learning using Entropy as Attribute Selection Measure", International Journal of Engineering and Technology, 2014.
12. (http://www.enggjournals.com/ijet/docs/IJET14-06-05-210.pdf)
13. Kishor Kumar Reddy C, Rupa C H and Vijaya Babu B, "A Pragmatic Methodology to Predict the Presence of Snow/No-Snow using Supervised Learning Methodologies", International Journal of Applied Engineering Research, 2014.
14. (http://www.ripublication.com/Volume/ijaerv9n21.htm)
15. Kishor Kumar Reddy C, Rupa C H and Vijaya Babu, "SPM: A Fast and Scalable Model for Predicting Snow/No-Snow", World Applied Sciences Journal, 2014.