Journal for Educators, Teachers and Trainers
The PROFESIOLab electronic, peer-reviewed, open-access Magazine

ISSN 1989-9572

DOI: 10.47750/JETT.2021.12.04.33

A Parametric Approach Using Active Shape Models for Facial Landmark Detection

K. Naresh1, T. Srujana2, D. Satheesh3

Journal for Educators, Teachers and Trainers, Vol. 12 (4)

https://jett.labosfor.com/

Date of reception: 15 Oct 2021

Date of revision: 04 Nov 2021

Date of acceptance: 05 Dec 2021

K. Naresh1, T. Srujana2, D. Satheesh3 (2021). A Parametric Approach Using Active Shape Models for Facial Landmark Detection. Journal for Educators, Teachers and Trainers, Vol. 12(4)251-260

Journal for Educators, Teachers and Trainers
The LabOSfor electronic, peer-reviewed, open-access Magazine

Journal for Educators, Teachers and Trainers
The LabOSfor electronic, peer-reviewed, open-access Magazine

JOURNAL FOR EDUCATORS, TEACHERS AND TRAINERS, VOL. 12 (4)
ISSN 1989 –9572
https://jett.labosfor.com/

A Parametric Approach Using Active Shape Models for Facial Landmark Detection

K. Naresh[1], T. Srujana[2], D. Satheesh[3]
[1]Department of Electronics and Communication Engineering
[1]Sree Dattha Institute of Engineering and Science, Sheriguda, Hyderabad, Telangana

ABSTRACT

On later tasks involving face recognition, three-dimensional face reconstruction, and other related tasks, facial landmark detection is essential. In this study, we introduce a parametric active shape model method for face landmark identification. In particular, we suggest using the active shape model (ASM) parameters to encode the landmark locations. Next, we estimate the ASM parameters by utilizing the power of cascade regression. The ASM parameters can be used to decode final landmark positions. Compared to previous similar work, this parametric method of predicting the landmark placements in another domain is more efficient and compact. We test the efficacy of our approach through experiments.

Keywords: Facial Landmark Detection, Active Shape Model (ASM), Parametric Modeling, Cascade Regression, 3D Face Reconstruction.

1. INTRODUCTION

Face is a complex multidimensional structure and needs good computing techniques for recognition. The face is our primary and first focus of attention in social life playing an important role in identity of individual. We can recognize a number of faces learned throughout our lifespan and identify that faces at a glance even after years. There may be variations in faces due to aging and distractions like beard, glasses or change of hairstyles. Facial features are extracted and implemented through algorithms which are efficient and some modifications are done to improve the existing algorithm models. Computers that detect and recognize faces could be applied to a wide variety of practical applications including criminal identification, security systems, identity verification etc. Face detection and recognition is used in many places nowadays, in websites hosting images and social networking sites. Face recognition and detection can be achieved using technologies related to computer science.

Features extracted from a face are processed and compared with similarly processed faces present in the database. If a face is recognized it is known or the system may show a similar face existing in database else it is unknown. In surveillance system if an unknown face appears more than one time then it is stored in database for further recognition. These steps are very useful in criminal identification. In general, face recognition techniques can be divided into two groups based on the face representation they use appearance-based, which uses holistic texture features and is applied in face image and feature-based, which uses geometric facial features (mouth, eyes, brows, cheeks etc.), and geometric relationships between them The proposed technique is based on coding and decoding of face images with emphasis on the significant of local and global features of face.

Face recognition is one of the major issues in biometric technology. It identifies and/or verifies a person by using 2D/3D physical characteristics of the face images. Facial landmarks are defined as the prominent features that play a discriminative role on the facial graphics, such as eye corners, nose tip, and mouth corners. Facial landmark detection is crucial for a series of tasks related to face, such as facial expression understanding, gaze estimation, and three-dimensional face reconstruction. The baseline method of face recognition system is the Eigen face by which the goal of the eigen face method is to project linearly the image space onto the feature space which has less dimensionality. One can reconstruct a face image by using only a few eigenvectors which correspond to the largest Eigen values, known as Eigen picture, Eigen face, elastic bunch graph matching and support vector machine. However, there are still many challenge problems in face recognition system such as facial expressions, pose variations, occlusion and illumination change.

Those variations dramatically degrade the performance of face recognition system. It is evident that illumination variation is the most impact of the changes in appearance of the face images because of its fluctuation by increasing or decreasing the intensities of face images due to shadow cast given by different light source direction. Therefore, the one of key success is to increase the robustness of face representation against these variations. In this proposed method the relevant information in a face image is feature extracted, encoded and then compared with a face database of models and then classified as known or unknown.

Face recognition has many challenges due to illumination variations, large dimensionality, uncontrolled environments, pose variations and aging. In the recent years, Face recognition get remarkable improvement and accuracy to overcome these challenges, but illumination change is still changing. The objective of this work goal is to implement the system (model) for a particular face and distinguish it from a large number of stored faces with some real-time variations as well. The proposed method is independent of any judgment of features like lighting problems pose variations eyes, different facial expressions images, with and without Glasses.

2. LITERATURE REVIEW

Active shape model (ASM) is an algorithm based on the Point Distribution Model (PDM). In the PDM, the geometric shapes of objects with similar appearances, such as human faces, human hands, hearts, lungs, etc., can be represented by the coordinate of several key points (landmarks) in series to form a shape vector. The active shape model treats facial landmarks as a deformable statistical model and then adjusts the parameters to fit the key points of the given facial image according to the detected local features. There are a lot of feature detection methods that can be used in the ASM, such as Boosted Haar Wavelets [4], Local Binary Patterns [5] and mutual information [6]. Regression techniques serve as the alternative to the above methods for local search. Unlike the conventional feature detection classifiers in ASM, which learn discrimination function between the local neighborhood and feature appearance, the regressors used to detect features learn the relationship between the local neighborhood feature and the displacement of the feature location ground truth. For example, Seise et al. [7] update feature locations with Relevance Vector Machine regressor. Wimmer et al. [8] regress local Haar Wavelet features to an objective function that can obtain the ground truth of the feature locations. Some studies have compared the performance of feature detection methods and regression techniques. By comparing Kernel Ridge Regression with a Bayesian Classifier approach, Everingham et al. [9] report that the simple classifier method has better performance in eye finding. Cristinacce et.al [10] show that the local feature regression model performs improved localization and is much more efficient. Some other related works focus on improving the shape model for better detection performance. In the original ASM, the face shape is represented through Principal Component Analysis (PCA). Zhou et al. [11] project the facial shapes into a tangent space and then use Bayesian inference to estimate both shape and pose parameters. Considering that the regressed shape is always the linear combination of all the training shapes, Cao et al. [12] constrain the shape model based on the linear combination instead of PCA. Since that ASM can obtain more accurate detection results at a faster speed and is also more robust with regard to illumination, activate shape model has become a suitable and efficient algorithm for facial landmarks often implemented on the mobile devices in recent years [13]. As for the cascade regression-based methods, they learn the regressor at each cascade level to iteratively update the landmark positions. The regressor at each cascade level maps the local features around the current key points to the landmark location ground truth. And therefore, the cascade regression-based methods vary in the input features and regressors. Both SDM [3] and LBF [14] use liner regressors at each cascade level. SDM directly uses SIFT features related to the face shape as the input features while LBF learns sparse binarization features in local areas based on the random forest regression model. Discriminative Response Map Fitting (DRMF) [15] gives a parametric model of the facial shape, using SVR as the regressor and HOG features as the input features. Recently, a cascade regression framework along with deep learning has achieved impressive performance. Deep Convolutional Neural Networks (DCNN) [16] combines coarse-to-fine cascade and geometric refinement to locate 68 facial landmarks. Instead of applying deep network directly, Coarse-to-Fine Auto-encoder Networks (CFAN) approach [17] cascade a series of consecutive stacked auto-encoder networks to infer the facial landmarks from the detected face region nonlinearly. An end-to-end deep convolutional cascade (DeCaFA) [18] architecture is introduced to incorporate the landmark-wise attention maps and intermediate supervisions into the deep cascade convolution network for landmark detection. To tackle the problem of landmark detection under occlusion, Wan et al. [19] propose to integrate a deep regression module and a deocclusion module into the cascade regression framework. However, the performance of deep learning-based methods highly relies on the scale of training samples. In addition, these methods are more likely to overfit the data. In this work, an active shape model parametric approach is proposed for facial landmark detection. We first encode the landmark locations by the ASM parameters. Then, we leverage the power of cascade regression to estimate the ASM parameters, which can decode the facial landmark locations finally. This active shape model parametric approach is more compact for the landmark location representation and can partially tackle the problem of overfitting.

3.PARAMETRIC APPROACH USING ACTIVE SHAPE MODELS FOR FACIAL LANDMARK DETECTION

We propose a new framework in the face recognition System by using Active Shape model (ASM). Initially we detect the face from the image. After that we extract the LBP feature. It is used to find the texture feature for the face image. The LBP operator assigned a label to every pixel of a gray level image. The label mapping to a pixel is affected by the relationship between this pixel and its eight neighbors. Active shape model (ASM) is a statistical model of the shape of objects which iteratively deform to fit to an example of the object in a new image. The shapes are constrained by the PDM (Point Distribution Model) Statistical Shape Model to vary only in ways seen in a training set of labeled examples. To locate a better position for each point one can look for strong edges, or a match to a statistical model of what is expected at the point. Then weighted matching will be applied between the input image and database images.

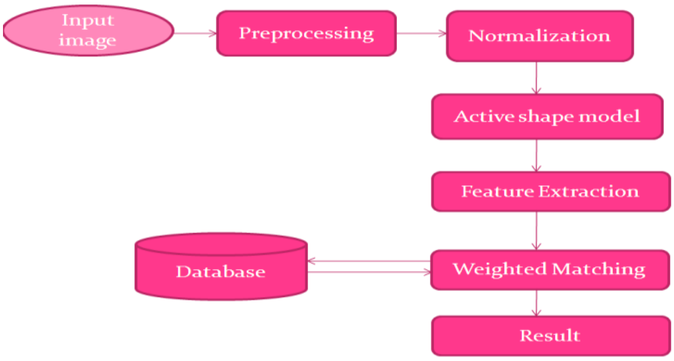


Figure 1: Proposed Model for Facial Landmark Detection

MODULES in the processing are given by

- Preprocessing
- Normalization
- Active shape model
- Feature Extraction
- Recognition

3.1 Module Description

Preprocessing

In noise removal process, initially we convert the image in gray. And then we filter the noise from the image. In Filtering we are applying Gaussian filtering to our input image. Gaussian filtering is often used to remove the noise from the image. Here we used wiener2 function to our input image. **Gaussian filter** is windowed filter of linear class, by its nature is weighted mean. Gaussian filter is named after famous scientist Carl Gauss because weights in the filter calculated according to Gaussian distribution.

Normalization

Normalization is a process that changes the range of pixel intensity values. Illumination changes caused by light sources at arbitrary positions and intensities contribute to a significant amount of variability. To address this issue, we present a new method for performing image normalization. The method used to remove shadows and specularities from images. All the shadowed regions are grayed out to a uniform color, eliminating soft shadows and specularities and hence creating an illumination invariant signature of the original image.

Active Shape model

Active shape models (ASMs) are statistical models of the shape of objects which iteratively deform to fit to an example of the object in a new image. The shapes are constrained by the PDM (point distribution model) Statistical shape model to vary only in ways seen in a training set of labelled examples. The shape of an object is represented by a set of points (controlled by the shape model). The ASM algorithm aims to match the model to a new image. It works by alternating the following steps:Look in the image around each point for a better position for that point. Update the model parameters to best match to these new found positions.

Feature Extraction

Initially we separate the image as patches. For each patch of image, we apply the LBP (Local Binary Pattern). The LBP operator assigned a label to every pixel of a gray level image. The label mapping to a pixel is affected by the relationship between this pixel and its eight neighbors of the pixel. If we set the gray level image is I, and Z0 is one pixel in this image. So we can define the operator as a function of Z0 and its neighbors, Z1, …, Z8. And it can be written as:

$$T = t\,(Z0, Z0\text{-}Z1, Z0\text{-}Z2, \ldots, Z0\text{-}Z8).$$

However, the LBP operator is not directly affected by the gray value of Z0, so we can redefine the function as following:

T = t (Z0-Z1, Z0-Z2, …, Z0-Z8).

To simplify the function and ignore the scaling of grey level, we use only the sign of each element instead of the exact value. So the operator function will become:

T = t (s(Z0-Z1), s(Z0-Z2), …, s(Z0-Z8)).

Where the s(.) is a binary function, defined as s(x) =1, x>=0; S(x) = 0, otherwise.

Histogram features

An image histogram is a type of histogram that acts as a graphical representation of tonal distribution describes the distribution of various bright and dark tones with in an image. During the scanning or image editing stage tones can be redistributed lightening a dark image (or) darkening a bright image. This histogram plots the no. of pixels for each tonal value. by looking at the histogram for a specific image a person will be able to judge the entire tonal distribution.

Image histograms are present on many modern digital cameras. The horizontal axis of the graph represents the tonal variations and the vertical axis represents the no.of pixels in that particular tone. For this histogram we are assuming a discrete function $h(r_k)=n_k$ Here rk is the kth gray level and n_k is the no. of pixels in the image at the gray level r_k.

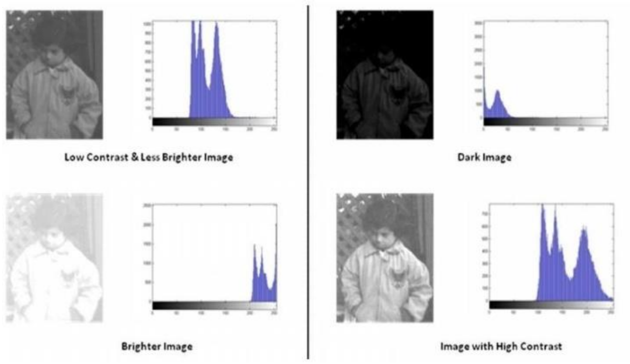


Figure 2: Different Types of Histogram Images with Different Contrasts

Recognition

Here the recognition process is identified by the weighted matching. The Euclidean distance for LBP based histogram features is computed for the test feature with the database features. The similarity is identified between the features. Finally identified image is displayed.

4. RESULTS& DISCUSSION

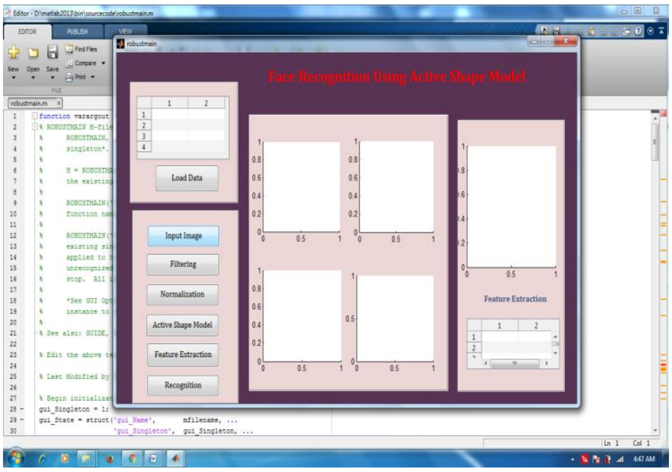
Data Set

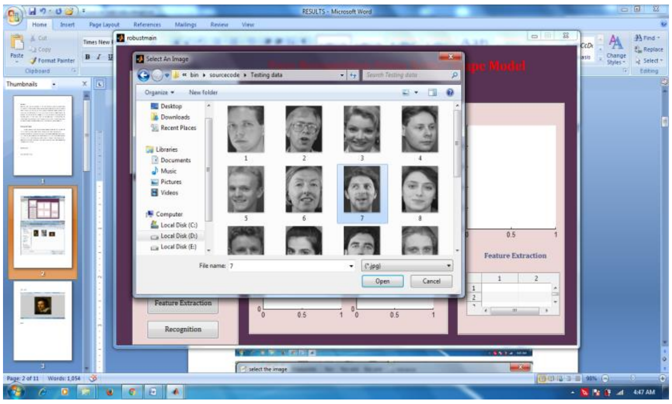
We have employed a set of images These images have been logically chosen by art historians in order to address different tasks such as (a) to test the relation of an unmediated image of the subject, e.g., a death mask to a work of portrait art like a painting, (b) to analyze a number of portraits of different sitters by the same artist to model artist's style, (c) to verify if the identity of the ambiguous subject in a given image is same as that of a known subject in a reference image.

Identity Verification

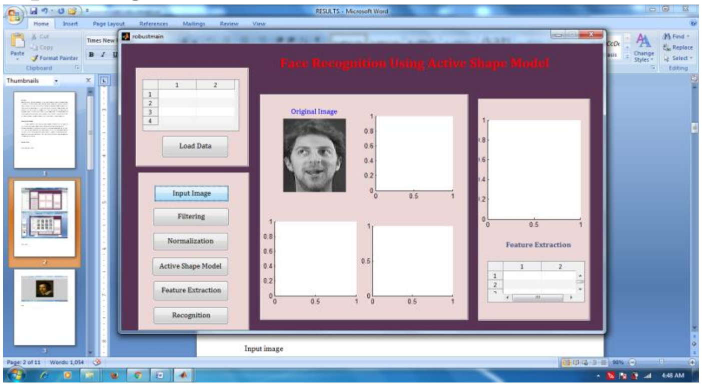
In order to examine the validity of the chosen approach, we consider similarity scores of the test image with artworks known to depict persons different from the one depicted in reference image. We call these images as distracters. In cases where enough works of the same artist is not available, we consider similar works of other artists. If a test image indeed represents the same sitter as in the reference image, not only should its score with the reference image be modelled by the match distribution, but also its scores with distracter faces should be modelled by the non-match distribution. The results of the proposed system are shown below.

Selection of input image

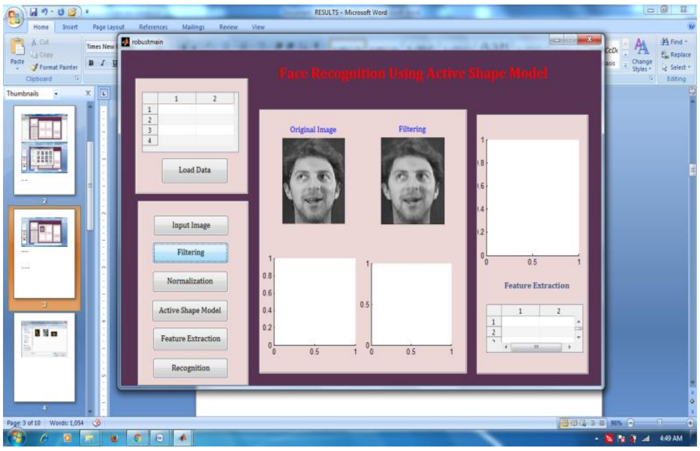




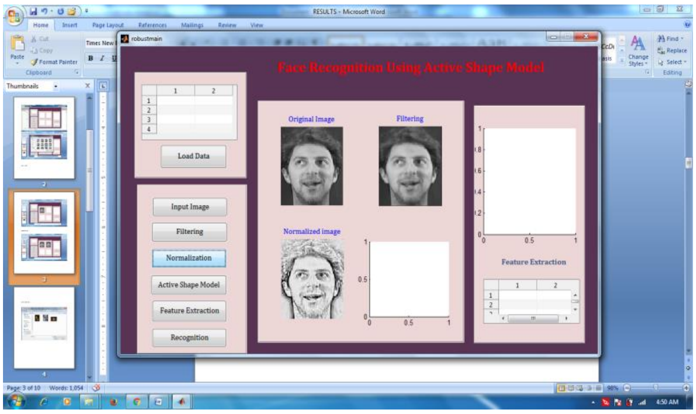
Input image



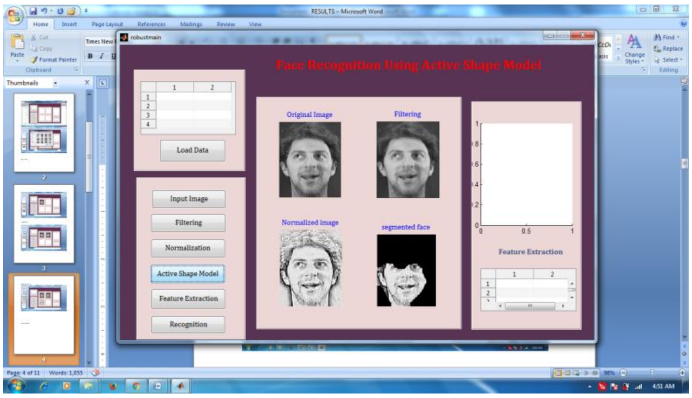
Filtering



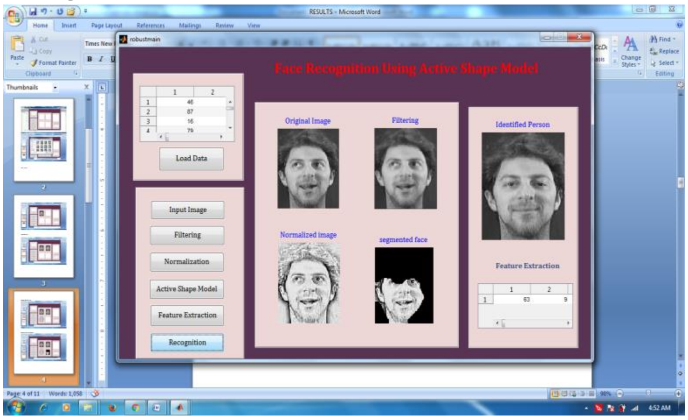
Normalization



Active Shape Model



Recognition



Advantages:

1. Security levels will be significantly improved.
2. The integration process is easy and flawless
3. High accuracy allows avoiding false identification
4. Facial Recognition System is fully automated
5. Time fraud will be excluded

Applications

1.Historical Persons Database
2.Politics
3.Education

4.Industrial
5.Security
6.Military

Conflict of interest statement
Authors declare that they do not have any conflict of interest.

REFERENCES

[1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, vol. 1. IEEE, 2001, pp. I–I.
[2] T. Ahonen, A. Hadid, and M. Pietikainen, "Face recognition with local ¨ binary patterns," in European conference on computer vision. Springer, 2004, pp. 469–481.
[3] N. D. Dowson and R. Bowden, "Simultaneous modeling and tracking (smat) of feature sets," in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 2. IEEE, 2005, pp. 99–105.
[4] M. Seise, S. J. McKenna, I. W. Ricketts, and C. A. Wigderowitz, "Learning active shape models for bifurcating contours," IEEE transactions on medical imaging, vol. 26, no. 5, pp. 666–677, 2007.
[5] M. Wimmer, F. Stulp, S. Tschechne, and B. Radig, "Learning robust objective functions for model fitting in image understanding applications." in BMVC, vol. 3, 2006, pp. 1159–1168.
[6] M. Everingham and A. Zisserman, "Regression and classification approaches to eye localization in face images," in 7th International Conference on Automatic Face and Gesture Recognition (FGR06). IEEE, 2006, pp. 441–446.
[7] D. Cristinacce and T. F. Cootes, "Boosted regression active shape models." in BMVC, vol. 2. Citeseer, 2007, pp. 880–889.
[8] Y. Zhou and H.-j. Zhang, "Bayesian tangent shape model: Estimating shape and pose via bayesian inference," in IEEE Conf. on CVPR. Citeseer, 2003.
[9] X. Cao, Y. Wei, F. Wen, and J. Sun, "Face alignment by explicit shape regression," International Journal of Computer Vision, vol. 107, no. 2, pp. 177–190, 2014.
[10] Y.-H. Lee, C. G. Kim, Y. Kim, and T. K. Whangbo, "Facial landmarks detection using improved active shape model on android platform," Multimedia Tools and Applications, vol. 74, no. 20, pp. 8821–8830, 2015.
[11] S. Ren, X. Cao, Y. Wei, and J. Sun, "Face alignment at 3000 fps via regressing local binary features," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1685–1692.
[12] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Robust discriminative response map fitting with constrained local models," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2013, pp. 3444–3451.
[13] E. Zhou, H. Fan, Z. Cao, Y. Jiang, and Q. Yin, "Extensive facial landmark localization with coarse-to-fine convolutional network cascade," in Proceedings of the IEEE international conference on computer vision workshops, 2013, pp. 386–391.
[14] J. Zhang, S. Shan, M. Kan, and X. Chen, "Coarse-to-fine auto-encoder networks (cfan) for real-time face alignment," in European conference on computer vision. Springer, 2014, pp. 1–16.
[15] A. Dapogny, K. Bailly, and M. Cord, "Decafa: deep convolutional cascade for face alignment in the wild," in Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 6893–6901.

[16] J. Wan, J. Li, Z. Lai, B. Du, and L. Zhang, "Robust face alignment by cascaded regression and de-occlusion," Neural Networks, vol. 123, pp. 261–272, 2020..