Journal of Information Systems Engineering and Management
2025, 10(11s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Age Estimation using Lightweight Convolution Neural Network

A.Annie Micheal [1,*],Augusta Kani G [2], Shanthi Thangam Manukumar[3],A.C.Santha Sheela[4]

[1] *Assistant Professor, Department of Computer science and Engineering, Sathyabama Institute of Science and Technology, Chennai,India, anniemichael.cse@sathyabama.ac.in*

[2] *Assistant Profesor, Department of Computer Science and Engineering, SRM Institute of Science and technology, Vadapalani, Chennai, India, augustag@srmist.edu.in*

[3] *Assistant Professor, Department of Computer science and Engineering, Sathyabama Institute of Science and Technology, Chennai,India, mshanthithangam@gmail.com*

[4] *Associate Professor, Department of Computer science and Engineering, Sathyabama Institute of Science and Technology, Chennai,India, santhasheela.cse@sathyabama.ac.in*

ARTICLEINFO

Received: 14 Nov 2024

Revised: 26 Dec 2024

Accepted: 10 Jan 2025

ABSTRACT

Introduction: Currently, age estimation is significant in sectors such as age-specific human-computer interaction, security control, surveillance monitoring, Electronic customer relationship management, and forensic art. Nowadays makeup is used by all groups of people. Makeup is now commonly used to boost an individual's confidence and self-esteem. Makeup can give a person a younger or older appearance depending on their makeup choice. This transformation is caused due to varied makeup products such as eyeliner, lipstick, foundation, compact powder, mascara, and eye pencil. These products impact the real face of the individual by concealing the age spots, scars, pimples, and wrinkles, highlighting the cheeks, and brightening the facial parts. Due to the transformation in the look, the apparent age varies from the real age. The estimation of real age is essential in various real-time applications.

Objectives: The application of makeup is a great challenge in estimating the age of the person in vision-based technology. This manuscript deals with deep learning-based age estimation with makeup.

Methods: The proposed framework is composed of data augmentation, face detection, face alignment, and age estimation. The data augmentation techniques used in our methodology are average blur, gaussian blur, histogram equalization, color jittering, bilateral filtering, detexturization, unsharp filtering, and gamma contrast. Multi-task Cascaded Convolutional Network (MTCNN) is adopted for face detection and alignment. Light CNN is used for age estimation with makeup.

Results: The methodology attains an MAE of 5.58 on the self-built Facial Makeup for Male and Female (FMMF) database.

Conclusions: This manuscript presents a deep learning based-age estimation with facial make-up and its results are also evaluated. The developed mythologies will improve the quality of applications such as human-computer interaction, surveillance systems, commercial development, content-based indexing, and searching, demographic studies systems, targeted advertising, and biometric system.

Keywords: Deep learning, Face detection, Face alignment, Multi-task Cascaded Convolution Network, Light CNN.

INTRODUCTION

In the current era, a person can increase their confidence and sense of self-worth by using makeup. Depending on the use of makeup, a person can appear younger or older with it. Makeup products like mascara, eye pencil, compact powder, foundation, lipstick, and eyeliner are to blame for this change. These products make the genuine face look better by lightening the areas of the face, bringing attention to the cheekbones, and hiding wrinkles, scars, and age spots. The apparent age differs from the actual age because of the change in appearance. The real age estimate is necessary for many real-time applications. Estimating an individual's true age without makeup is a difficult task in computer vision. It's difficult to determine someone's real age when they wear makeup.

LITERATURE REVIEW

Cunjian Chen et al. [1] examined the impact of cosmetics on automated gender and age prediction systems. A person's perceived age can be altered with makeup. This is accomplished by using a light foundation and concealer to cover wrinkles and age spots, concealer and powder to brighten shadows created by wrinkles around the mouth, nose, and eyes, and highlighter and blush to color and highlight the cheeks. This study aims to investigate the impact of makeup on an automated age estimation system. Cropped and aligned faces were used to reduce additional confounding factors (such as hair and accessories). The age of an individual was estimated using the OpenBR software, and then compute the variations in estimated age before and after makeup for each individual. Vairavanet al [2] estimated the age of an individual with facial makeup. A method for face detection was initially used to normalize the facial image. The unique features such as shape, texture, and regions were extracted. Local Gabor XOR pattern was used for the extraction of texture features. Active Appearance Model was employed to extract the shape feature. The eye portion was separated from the detected face for the region area. Age estimation was done using an Artificial Neural Network (ANN) and achieved an MAE of 6.1 on a self-built database. Mousavi S.M.H and Lyashenko V [3] used color-depth images to estimate age with makeup. Using depth sensor, it is possible to create 3- dimensional model of the face and extracts wrinkles that were concealed under makeup. This feature aids in the age estimation procedure by assessing the number of wrinkles. The face was detected using Viola and Jones algorithm. In order to obtain depth images, the standard deviation filter [4] and elliptical fitting [5] techniques were applied after determining the closest sample from the subject to the sensor that is noisy tip. The next step is to sum the grey values of both color and depth images, as well as normalization, was performed between the youngest and eldest age in the dataset. The performance of the system was evaluated with FG-NET [6], IIITD Kinect RGBD, Face Grabber, MIFS, AGFW, and Depth Makeup Dataset (IDMD) datasets. Only images with makeup were chosen from FG-NET, IIITD Kinect RGBD, and Face Grabber datasets because not all of the subjects were in makeup. The MAE for each dataset in four different ranks was calculated. The rank represents the number of samples used in each dataset. The system has a high degree of speed but its accuracy is low. The major disadvantage of this paper is the experiment is performed with only female subjects. RanranFreng et al. [7] carried out a thorough statistical analysis of the impact of makeup on facial features like the skin, eyes, and lips. Based on an input face image, the statistical study suggests a technique for determining whether or not makeup is applied. The Young Index (YI), which is used to estimate the age of female subjects, is then used to quantify the makeup effect. When the makeup effect is considered, the method proposed in this research can improve accuracy by 0.26-9.76 percent in MAE and 0.9-6.7 percent in CS (Cumulative Score) when compared to other age estimating methods. Only Limited research is available for age estimation with facial makeup.

PROPOSED METHODS

The framework for age estimation with makeup is composed of data augmentation, Face Detection and Alignment, and age estimation as depicted in Figure 1. Data augmentation increases the dataset. The face detection module detects the face in an image. The face alignment module, aligns the face based on the detected landmark. Shape information plays a predominant role in estimating age with makeup. Light-CNN focuses on shape cues, hence it is adopted for age estimation with facial makeup.

A. Data Augmentation

Data augmentation is a technique of increasing the diversity and amount of data by applying transformation methods to the existing dataset. Data augmentation solves the overfitting problem and improves the generalization of the deep neural networks. The data augmentation technique increases the number of data corresponding to each sample without changing the image fidelity and visual quality of the images. Henceforth, the augmented images are used by the CNN model during the training stage to enhance the learning parameters of CNN architecture and attain the best model for classification or recognition problems. The data augmentation techniques used in our methodology are Histogram equalization, Color jittering, Bilateral filtering, Unsharp filtering, and detexturization.

1.Average blur

By averaging all the pixels in the surrounding of a central pixel, average blur substitutes the average for the central pixel.

2.Gaussian blur

Gaussian blur an image using a Gaussian function.

3.Histogram equalization

Histogram equalization [8] is used to enhance the contrast of a face image by spreading out the intensity range of the image. As a result, regions of lower contrast will gain higher contrast and the histogram of the output image is equally distributed.

4.Color jittering

Color jittering methods include random changes of the contrast, hue, brightness, and saturation of the image.

5.Bilateral filtering

Bilateral filter [9] is a non-linear filter. This filtering technique reduces the noise and smooths the images while preserving the sharp edges.

6.Detexturization

The face images are detexturized using an anisotropic diffusion approach [10]. This approach removes the texture by preserving smooth homogenous regions and edges. Detexturized face images will allow light CNN to learn discriminative features even in the presence of foundation used to cover scars, open pores, and moles.

7.Unsharp filtering

Unsharp filtering is used for enhancing the edges in the face image. The resulting edge map will enable the light CNN to learn facial lines to be matched with images with facial contouring like giving shape to facial components such as lips and nose bridge.

8.Gamma contrast

Gamma contrast improves the image luminance by enhancing the quality of the image in dark or blurry areas. Figure 2 shows the sample output of the data augmentation techniques.

B.Face Detection and Face alignment

For face detection and face alignment, Multi-task Cascaded Convolutional Networks (MTCNN) [11] is employed. MTCNN is divided into three phases. Proposal Network (P-Net) is utilized in the first stage to obtain the bounding box regression vectors of the candidate windows. To calibrate the candidates, the estimated bounding box regression is employed. It uses non-maximum suppression (NMS) to combine highly overlapped candidates. All candidates are fed into Refine Network (R-Net) in the second stage, where a significant portion of non-faces candidates are rejected. Moreover, it carries out NMS candidate merge and calibration bounding box regression. In the third sage, the output network (O-Net) produces the five facial landmark positions and the final bounding box.

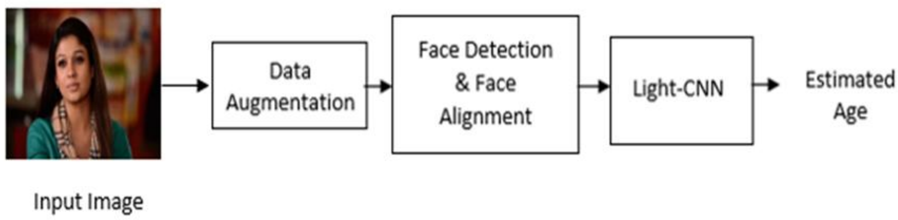

Figure 1: The architecture of Age estimation with facial makeup

C.Face classification

The classification of faces is regarded as a two-class problem. The cross-entropy loss for each image x_i is defined as

$$L_i^{\det} = -(y_i^{\det}\log(p_i) + (1-y_i^{\det})(1-\log(p_i))) \tag{1}$$

where p_i denotes the probability that the image is a face and $y_i^{\det} \in \{0,1\}$ represents the ground-truth label.

D.Bounding box regression

We predict the offset between each candidate window and the closest ground truth, which consists of the left top, height, and width of the bounding boxes. For each image, the Euclidean loss is denoted as

$$L_i^{box} = \left\| \hat{y}_i^{box} - y_i^{box} \right\|_2^2 \tag{2}$$

where $\hat{y}_i^{box}$ regression targets and y_i^{box} denotes the ground-truth coordinate.

E.Facial landmark localization

The problem of facial landmark localization is formulated as a regression, and the Euclidean loss is minimized as follows

$$L_i^{landmark} = \left\| \hat{y}_i^{landmark} - \hat{y}_i^{landmark} \right\|_2^2 \tag{3}$$

where $\hat{y}_i^{landmark}$ denotes the coordinate of the face landmark and $y_i^{landmark}$ denotes the ground-truth coordinate. The faces are aligned using the five recognized facial landmarks.

(a) Average Blur (b) Gaussian blur (c) Histogram equalization (d) Color jittering

(e) Bilateral filtering (f) Detexturization (g) Unsharp filtering (h) Gamma contrast

Figure 2: Sample output for data augmentation techniques

F.Light CNN

The detected face is fed into Light CNN. The Light CNN obtains a deep face representation by acting as feature extractors. Light CNN is composed of four max-pooling layers, five convolution layers, four Network In Networks (NINs), and the Max-Feature-Map Activation function. The structure of Light CNN is shown in Figure 3. Max-Feature-Map (MFM) activation function is used for feature selection and compact representation. During backpropogation, MFM selects the best feature at each location, which produces a binary gradient. Because of MFM's gradient, stochastic gradient descent (SGD) can only affect response variable neurons in the training stage. During testing stage, MFM can activate the maximum of two feature maps to attain more competitive nodes from previous convolution layers. The activation of small number of neurons are suppressed when MFM reaches the max function.

G.Max-feature-Map activation Function

An input convolution layer $x^n \in R^{H \times W}$,where n={1,2,...,2N}, H denotes the spatial height and W denotes spatial weight of the feature map. The MFM activation function combines two feature maps and produces element-wise maximum one as follows

$$\hat{x}_{ij}^{k} = \max(x_{ij}^{k}, x_{ij}^{k+N}) \tag{4}$$

The gradient of Eq (4) can be written as

$$\partial\hat{x}_{ij}^{k} = \begin{cases} 1, if x_{ij}^{k} \geq x_{ij}^{k+N} \\ 0, otherwise \end{cases} \tag{5}$$

The pre-trained light CNN has an output softmax layer set to the number of categories in the training stage. Rather than using a set of distinct classes to estimate an individual's age, we would prefer to use a continuous value. Consequently, in our work, age estimation is regarded as a regression problem. A 1-layer neural network is used in place of the last softmax layer in order to learn the age regression function. Equation (6) provides the mean square error (MSE) loss function, which is optimized to learn the regression.

$$L = \frac{1}{N}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{6}$$

where L denote the average loss for all the training samples, $\hat{Y}_i$ denote the estimated age, and Y_i denote the real age.

EXPERIMENTAL RESULTS

Existing makeup datasets like Makeup Induced Face Spoofing (MIFS) [12]. YouTube Makeup Database (YMU), Facial cosmetic database (FCD) [13], Virtual makeup dataset (VMU) [14], and Makeup in the wild (MIW) [15] comprises only female images. Therefore, Facila makeup for male and female (FMMF) database has been generated which comprises actor and actress images. FMMF database comprises fo 3840 images (1734 male and 2106 female) with 640 subjects. Each subject has 6 images (3 images with makeup and 3 images without makeup). Figure 4 shows the sample images from the FMMF database. Data augmentation is performed with various techniques to increase the FFMF dataset. The data augmentation enables CNN architecture to perform better during fine-tuning and makes the classification more effective and robust. Face alignment and detection are handled by multi-task cascaded CNN. Four types of data annotation are used in face detection and alignment module training, which are as follows:

(i) Negatives: Areas where any ground-truth faces have an Intersection-over-Union (IoU) ratio of less than 0.3

(ii) Positive: IoU greater than 0.65 to a face of ground truth

(iii) Part faces: IoU between 0.4 and 0.65 to a ground truth face.

(iv) Landmark faces: faces labeled with the positions of five landmarks. Bounding box regression uses part faces and positives, face classification tasks use negatives and positives, and facial landmark localization uses landmark faces.

Using positives, negatives, part faces, and landmark faces from the FMMF dataset, the P-Net, R-Net, and O-Net models of the face detection and alignment module are refined. The pre-trained model on WIDER FACE [16] and CelebA [17] initializes the models. The five identified facial landmarks are used to align the face images from the detected faces. After being identified, the face is resized to 256 × 256 and fed into a light CNN for training. Next, using the FFMF dataset that the pre-trained model on CASIAWebFace initialized, light CNN is adjusted. Table 1 shows the MAE of different CNN architecture for the original and aligned image. Figure 5 shows the comparison of different CNN architectures. Figure 6 shows the sample output for age estimation with facial makeup.

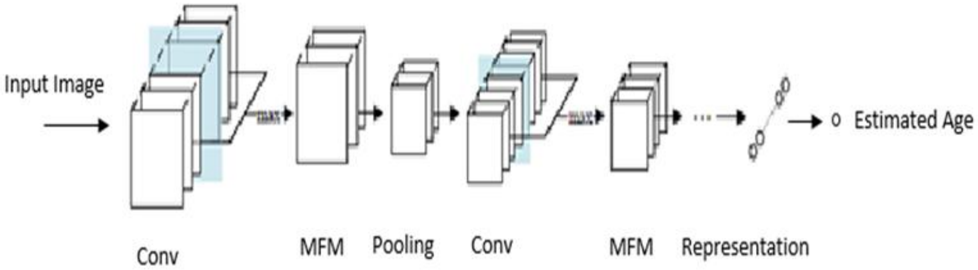

Figure 3: The framework of Light CNN

(a) Sample images without facial makeup

(b) Sample images with facial makeup

Figure 4: Sample images from the FMMD database.

Table 1:MAE of different CNN architecture

CNN	Original	Aligned
WebFace	13.26	10.02
FaceNet	11.86	8.74
VGG	13.64	10.43
Light CNN-4	9.54	7.05
Light CNN-9	7.43	**5.58**

Figure 6. Age estimation with makeup output

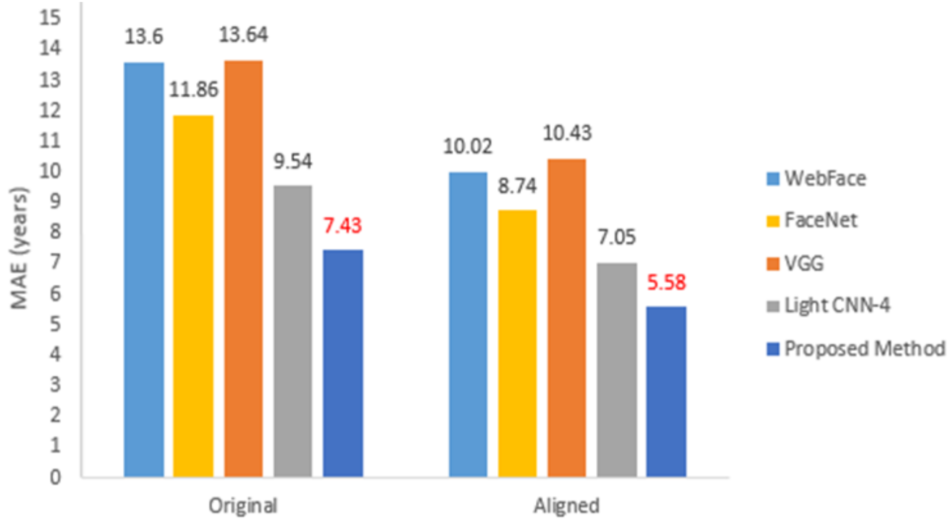

Figure 5: Comparison for different CNN

CONCLUSION

Recently usage of makeup is high and hence age estimation with makeup is quintessential. This framework for age estimation with makeup is composed of data augmentation, face detection and alignment, feature extraction, and age estimation. The data augmentation techniques used in our methodology are an average blur, gaussian blur, histogram equalization, color jittering, bilateral filtering, detexturization, unsharp filtering, and gamma contrast. Multi-task Cascaded Convolutional Network (MTCNN) is used for face detection and face alignment. The Light CNN learns a deep face representation as feature extractors. The methodology attains an MAE of 5.58 on the self-built Facial Makeup for Male and Female (FMMF) database. The developed mythologies will improve the quality of applications such as human-computer interaction, surveillance systems, commercial development, content-based indexing, and searching, demographic studies systems, targeted advertising, and biometric system.

REFRENCES

[1] Chen, C.; Antitza, D.; Arun, R. Impact of facial cosmetics on automatic gender and age estimation algorithms. International Conference on Computer Vision Theory and Applications. 2014, 182-190.

[2] Vairavan, T.; Vani. An Efficient Age Estimation System with Facial Makeover Images Based on Key Points Selection. International Arab Journal of Information Technology. 2017, 14, 8-17.

[3] Mousavi.; Seyed Muhammad, H.; VyacheslavLyashenko. Age Estimation in Make-up (Cosmetics) Using Color and Depth Images. 2019.

[4] Chang, D.C.; Wu,W.R. Image contrast enhancement based on a histogram transformation of local standard deviation. IEEE transactions on medical imaging.1998, 17, 518-531.

[5] Gander, W.; Golub, G.H; Strebel,R. Least-squares fitting of circles and ellipses. BIT Numerical Mathematics. 1994,34, 558-578.

[6] Micheal, A. A.; Barathi, B. A.; Micheal, A. A. A Framework for Age Estimation Using Facial Features: A Comprehensive Survey. Navigating Usability and User Experience in a Multi-Platform World. 2024, 249-278.

[7] Feng.; Ranran.; Balakrishnan, P. Quantifying the makeup effect in female faces and its applications for age estimation. 2012 IEEE International Symposium on Multimedia. 2012, 108-115.

[8] Gonzalez, R.C.; Woods, R.E.; Digital image processing, 3rd edn. 2007.

[9] Banterle, F.; Corsini, M.; Cignoni, P.; Scopigno, R. A Low-Memory, Straightforward and Fast Bilateral Filter Through Subsampling in Spatial Domain. Computer Graphics Forum.2011,31, 19–32.

[10] Perona, P.; Malik, J. Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence.1990, 12, 629–639.

[11] Zhang, K.; Zhang, Z.; Li, Z. Qiao Y, Joint face detection and alignment using multitask cascaded convolutional networks. IEEE signal processing letters.2016, 23,1499-1503.

[12] Chen, C.; Dantcheva, A.; Swearingen,T.; Ross, A. Spoofing Faces Using Makeup: An Investigative Study. IEEE International Conference on Identity, Security and Behavior Analysis.2017, 1-8.

[13] Eckert, M.; Kose, N.; Dugelay, J. Facial Cosmetics Database and Impact Analysis on Automatic Face Recognition. IEEE 15th International Workshop on Multimedia Signal Processing (MMSP). 2013, 434-439.
[14] Dantcheva, A.; Chen, C.; Ross, A.; Wv, M. Can Facial Cosmetics Affect the Matching Accuracy of Face Recognition Systems ?. IEEE Fifth international conference on biometrics: theory, applications and systems (BTAS).2012, 391-398.
[15] Chen, C.; Dantcheva, A.; Ross, A. Automatic Facial Makeup Detection with Application in Face Recognition. International Conference on Biometrics (ICB).2013, 1-8.
[16] Yang.; Shuo.; Ping,L.; Chen-Change, L.; Xiaoou,T. Wider face: A face detection benchmark. in Proceedings of the IEEE conference on computer vision and pattern recognition.2016, 5525-5533.
[17] Liu, Z.; Luo, P.; Wang, X.; Tang,X. Deep learning face attributes in the wild. in IEEE International Conference on Computer Vision. 2015, 3730-3738.