Age and Gender Recognition With Facial Images Using CNN

D Ramya
Department Of Information Technology
Vardhaman College of Engineering
Kacharam, Hyderabad, India
danduramya23@gmail.com

P R Harshith
Department Of Information Technology
Vardhaman College of Engineering
Kacharam, Hyderabad, India
prharshith37@gmail.com

P V Sai Sumeeth
Department Of Information Technology
Vardhaman College of Engineering
Kacharam, Hyderabad, India
saisreenath11@gmail.com

K Santhosh Kumar
Department Of Information Technology
Vardhaman College of Engineering
Kacharam, Hyderabad, India
santhoshkumar1208@gmail.com

***Abstract*—In recent times lot of efforts have been put on age and gender prediction. Age and gender can be measured accurately under certain conditions like stationary lighting conditions, front faces. Type of camera, positioning, and lighting conditions have huge effect on accuracy. In previous papers age and gender prediction is done using datasets or saved images. In this paper we proposed AGE AND GENDER RECOGNITION WITH FACIAL IMAGES , the goal is to capture live image of person and estimate age and predict gender of person using OpenCV Deep Learning. Image processing is used to enhance images captured by cameras, satellites, planes and cameras. Images are processed using OpenCV various methods and calculation based analysis. The available models fall short of required accuracy level to use these models in real-world applications due to significant intra-class variance of face images such as illumination, pose, scale change. In this paper Haar Cascade Classifier is being used to extract facial features which is the most important steps of the model.**

***Keywords*—** Gender prediction, Age estimation, OpenCV framework, Convolutional Neural Network, DNN Face Detector

I. Introduction

Age detection is the process of getting to know about a person's age with the usage of a still photo or live image of their face. A facial image is the source of a variety of significant pieces of information, including identification, expression, emotion, gender, race, age, etc. Wide interest has recently grown in the topic of computer vision, especially in face identification, detection and localization of facial landmarks. Age estimation is an automatic method of categorizing an image of a face according to an exact age or age range. Over the past few years, many works have been offered to predict age and gender. Previous research mostly relied on manually derived facial photographs that were then classified. However, due to the tremendous achievement of getting to know models in numerous computer imaginative and prescient problems over the past decade, recent work on age and gender prediction has mostly shifted to deep neural network-based models. In this project we worked on OpenCV, Deep Learning to estimate age and gender from live images.Deep neural networks are one of the main types of neural networks used for image classification and image recognition. Convolutional neural networks were first established in 1995 by Yann LeCun and Yoshua Bengio in order to correctly classify images and recognise handwritten control numbers. The structure of LeCun's CNN is made up of a series of different layers S-layers and C-layers, also known as convolution and sub sampling, with one or more fully connected layers at the ends. Convolutional Neural Networks are a unique kind of multi-layer neural networks; in terms of their shape and structure, they resemble their progenitor. The facial photos are classified based on age and gender via the feature extraction procedure, which extracts a feature matching to each. With a powerful picture pre-processing technique, we specifically address the significant differences in unfiltered real-world faces. This approach prepares and analyses such facial images before supplying them to the CNN model. In section III(H) we discuss the analysis of results of the approach used. It is followed by conclusion in section IV.

II. Literature Review

In this research [1], they suggested a hierarchical method for calculating an individual's age automatically and used component-based representations to study how ageing affects various face features. The performance of the suggested automatic age estimation method is compared to the evaluation of human perceptual ability to estimate age using crowdsourced data from the Amazon Mechanical Turk service. Eyes and the nose are more age-predictive than other face features, according to experimental findings from the FG-NET, MORPH Album2, and PCSO databases (forehead, eyebrows, mouth, and shape).

This paper [2] discusses To solve the issue of age estimation and gender classification, they have proposed two methods in this work. Learning transfer using pre-trained models and unique architectures. In order to avoid overfitting, these pre-trained models were quite helpful. When tested on actual photos, we discovered that the model generalises very effectively and with little overfitting. They intend to extend our research to larger and more balanced datasets, allowing us to conduct additional experiments to examine biases and enhance the generalizability of our approach. They intend to build on this outcome and the deep learning community in their subsequent work by using it as a springboard for additional improvisation and creativity.

The Deep EXPectation (DEX) method described in this paper [3] employs a CNN that has been pretrained on ImageNet. Additionally, we indexed age-related face photos from the internet to compile the largest publicly available dataset of its sort to date for pre-training the CNN. Additionally, age-tagged facial photos are used to fine-tune our CNN. We demonstrate an extention over problem of the CNN by presenting the age estimation problem 20 networks' predictions from cropped face photos are combined using DEX.

This paper [4] discusses that it has more than 9,000 IDs and between 80 and 800 photos per ID, for a total of more than 3 million photographs. Additionally, a pipeline is suggested to create a dataset that encourages age and postural variety within each individual. Filtering is used to reduce label noise. Third, we offer template comments on the test set so that users may directly assess how well posture and age detection work. They suggested a collection strategy to gather the VGGFace2 high-quality dataset for a variety of positions and ages. We also demonstrate that deep learning models are trained on VGGFace2 attain state-of-the-art performance.

In order to evaluate the planned work in this research, several experiments were carried out [5]. The specifications of the benchmark database are covered in this part along with the characteristics of the experiments that were carried out. The Adience benchmark is among the most popular datasets developed for age estimation from facial images. An audience evaluates the suggested work's effectiveness in this project. It consists of innumerable facial images uploaded to the Flickr website by 2284 users in 8 different age categories. Table III shows the number of images for each age category in addition to the distinction between men's and women's age categories.

This article [6] explores Deep learning was employed in this study to determine the actual age with still images. Using his VGG-16 architecture, they have already received pre-training in deep learning for categorising images [13]. Due to the lack of age-apparent images, we also investigated the benefits of modifying face images from the internet that had been crawled by accessible age. He looked through 500,000 publicly accessible images of famous people from IMDB as well as Wikipedia. The largest body of age prediction data to date has been made public. In contrast to training the CNN with straight regression, they express the age challenge as a depth categorization job and then improve it by employing the softmax mean.

This paper [7] explores The suggested method uses a fully programmable model with a multiple processors that supports SIMD functionalities together with an inter processor communication scheme. The dynamic data architecture methodology was chosen as a high-level programming solution because it enables parallelism, scale, modularity, and data-driven features. Each of them complies with the requirements for vision system design. Using a data flow model, algorithms can be represented as actors in a directed graph where results denote processing and arcs denote data transfer. When a data flow is handled by the availability of the data itself, it is said to be "data-driven."

This paper [8] explores In this white paper, new representations, methods, and insights are listed. These are quite common, employ a variety of techniques, and have potential applications in image processing and computer vision that are more comprehensive. Last but not least, this article covers several exhaustive experiments on a challenging facial recognition software dataset which has been the focus of extensive research. Faces are represented in this dataset under a range of lighting, scale, posture, and camera variation conditions. Conducting trials with such vast and complex data sets is difficult and time-consuming. However, it is unlikely that systems running in this environment will be delicate or restricted to a certain set of situations. Furthermore, it's unlikely that experimental artefacts contributed to the conclusions drawn from this dataset.

This methodology for precise age estimation methods in this work can be further developed [9]. Results might be improved by using several samples from various geographic regions. The findings demonstrate that children may be automatically identified using their fingerprints. Given that computers used by internet users now employ fingerprint scanners as a type of biometric password, this contemporary technology can be improved to recognise minors online. As a result, you can use this method to reduce the risks your child faces online.

In this paper [10], In this article, we explored the application of neural networks to estimate human age. The uniqueness of our research project is that we consider age groups more finely tuned than most previous studies and apply our experiments to real human face images. After collecting real face images, facial features were extracted and used as input for training a multi-layer perceptron neural network (MLP). The results model an accurate age estimator that has minimal estimation error for MLP networks and can be used in many useful applications, such as age-based access control and age-adapted human-machine interaction. It shows that it can be considered an excellent method.

III. Proposed Method

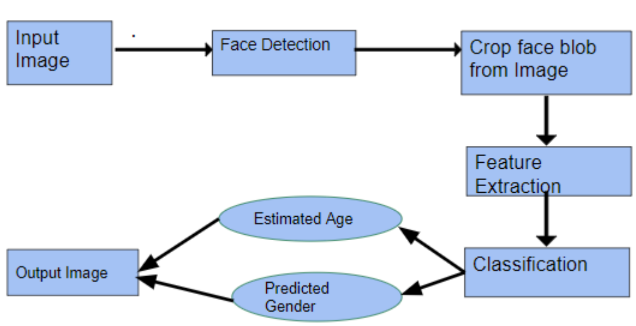

Fig. 1. Flow Chart Representation

A. Face Detection

Face detection is a method for locating or identifying the face in digital photographs or video frames. A human can easily and rapidly recognize the faces. For us, it is a simple task, but for a computer, it is challenging. OpenCV is a large open-source library for image processing, machine learning, and computer vision. It now plays a significant part in real-time operation, which is crucial in modern systems. It processes photos and videos to recognise objects, faces, and even human handwriting Face recognition. This model uses caffe models for image recognition and then classifying them based on age range and gender to which they belong to. A special case of object class recognition is face recognition.A deep learning framework called Caffe enables users to build picture classification and segmentation models. User models are initially created and saved as plain text files known as PROTOTXT files. The user's trained models are saved as a CAFFEMODEL file after the user uses Caffe to train and improve their model. Files called CAFFEMODEL are binary protocol buffer files. So, unlike PROTOTXT files, they cannot be opened, checked out, and edited in a source code editor. The purpose of CAFFEMODEL files is to enable the integration of Caffe-trained image classification and image segmentation models into applications.

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)} \qquad (Eqn.1)$$

B. Creating DNN For Models

Input, output, and a few hidden layers are all present in deep neural networks. These networks are capable of handling non-linearity as well as unstructured and unlabeled data. In OpenCV, DNN is being created in order to load pre-trained models and pass it to model files. For that we first declare variables in which file paths of model files are stored. These vary from model to model, I have taken values from caffe model that is used in the project. In order to develop DNN models we first need to define layers of DNN. The next step is training the model. After training the model, trained model is used to make predictions.

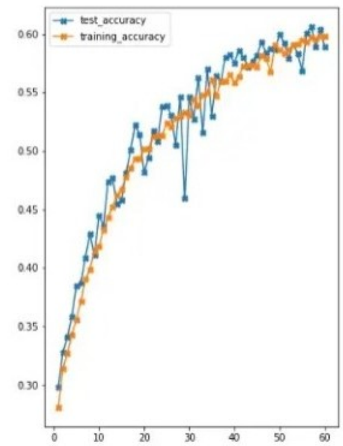

Fig. 2. Simulation Graph Representation

C. Crop Face Blob From Image

In this step, For the input image, a 4-dimensional array or blob is returned. Additionally, for further pre-processing you can use the cropped image so that it complies with your input criteria. To change your image, you can use its many attributes. This model uses DNN face detector for image recognition. Certain part of face or image is bolbed and then feature extraction is done. The caffemodel and prototxt files are used if we wish to employ Caffe's floating point model. If not, we employ the TensorFlow model.

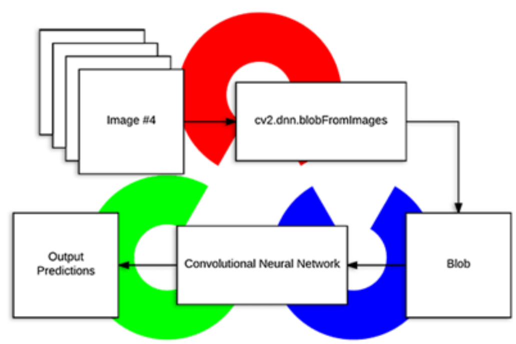

Fig. 3. General structure of the proposed system

D. FEATURE EXTRACTION

All of the information we gather is large in real life. We require a process to comprehend this data. Processing them manually is not possible. This is where the idea of feature extraction enters the picture. Dimensionality reduction techniques that divide and compress the initial set of uncooked facts into smaller, extra possible agencies encompass characteristic extraction. This makes processing less difficult. The truth that those huge datasets contain a extensive sort of variables is the most vital feature, Processing those variables calls for a number of computational power. To efficiently lessen the quantity of information, function extraction allows extract the pleasant functions from those massive amounts of statistics by means of selecting variables and combining them into functions. those functions are easy to use and as it should be and simply describe the real statistics set. To facilitate pattern identification, function extraction is a unique sort of dimensionality reduction utilized in pattern recognition and picture processing. locating the maximum relevant records inside the true facts and representing it in a low-dimensional area is the essential cause of feature extraction. DNN face detector in OpenCV is used for this process

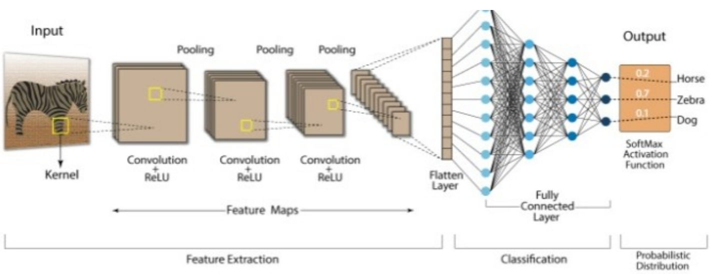

Fig. 4. Feature Extraction Process

E. GROUPING IMAGE INTO BINS

Binning groups images into bins, with each bin representing a specific age range. Continuous data can be classified into exact ranges using binning. Shoes are a great example of binning because they have much more specific foot measurements than shoe sizes available in stores.Binding every pixel to a color histogram usually determines the color composition of an image. This is the first step in image processing when This obviously results in less image jitter, but image comparison has some important advantages. Various combinations were attempted with the aim of recreating typical aging periods that individuals resembled as closely as possible. For the younger age groups in the data, several years are included in each class to ensure that individuals in each class are as similar as possible. The result was using the following bins: ['(0-2)', '(4-6)', '(8-12)', '(15-20)', '(25-32)', ' (38-43)', '(48-53)', '(60-100)'] correspond to 8 different classes.

$$Number\,of\,bins = 3 + log_2(n) * log(n) \qquad (Eqn.3)$$

F. CLASSIFICATION

Convolutional neural networks are a class of neural networks mainly designed for processing gridded records inclusive of time series and pictures. CNNs are commonly composed of convolutional layers and absolutely related layers. each convolutional layer performs characteristic extraction, and the FC layer uses the results from the convolutional layers to classify images into particular labels.

$$Covolution\ Z = X * f \qquad (Eqn.4)$$

A CNN is a device studying algorithm that allows a computer to expect and obtain photo functions and determine if the name of a newly captured picture must be derived from the ones capabilities. Deep gaining knowledge of uses emerging algorithms and artificial neural networks to educate machines and computer systems to examine from revel in, classify facts, and apprehend statistics and photos as the human brain does.in order CNNs are a class of artificial neural networks used in deep learning and are frequently used for object and image recognition and classification. CNNs are widely utilised in a variety of jobs and activities, including speaker identification with natural language processing video analytics, image processing, computer vision, and obstacle identification in self-driving cars. owing to their substantial contribution to these quickly developing and increasing domains, CNNs are widely used in deep studying.

$$Dimension\,of\,image = (n, n) Dimension\,of\,filter = (f, f)$$

$$Dimension\,of\,output = ((n - f + 1), (n - f + 1)) \quad (Eqn.2)$$

G. PREDICTION OF GENDER AND APPROXIMATED AGE

The gender and expected ages of the supplied image are shown below. Two alternative models were used to predict age and gender. Age deploy as well as gender deployment models are used to anticipate the age and sexual orientation of the input image. Eight different age lists are mentioned. Given an input image, ages are predicted in the following range: ['(0-2)', '(4-6)', '(8-12)', '(15-20) ', '(25-32)", "(38-43)", "(48-53)", "(60-100)"], gender is two attributes assigned by age list ["Male" ,"female"].

H. Result Analysis

a) Comparing The Outputs: The output age generated by machine vs real age

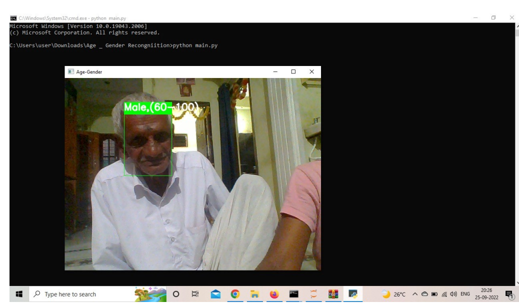

Fig. 5.

TABLE I
OUTPUT COMPARISION TABLE

Image	Actual Ages	Predicted Ages
Fig. 5	72	item 60-100
Fig. 6	21	item 21-32
Fig. 7	21	item 21-32
Fig. 8	5	item 04-06

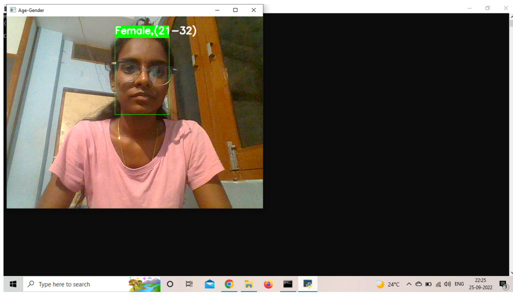

Fig. 6.

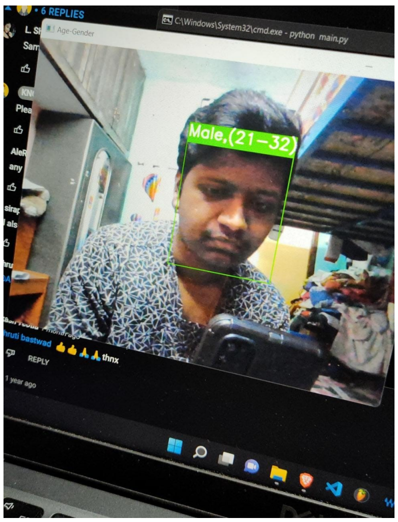

Fig. 7.

IV. Conclusion

Although age and gender categorization issues have been addressed by a variety of prior techniques, most of this research has, up until recently, been on limited photos captured in lab settings. These settings don't accurately capture the look variances found in real-world pictures found on social media platforms and online archives. Internet photos, however, are not only more difficult, but more plentiful. The study concludes that research on age and gender has become an interesting topic in recent times. Many strategies in the past have focused on issues of age and sexuality, but more

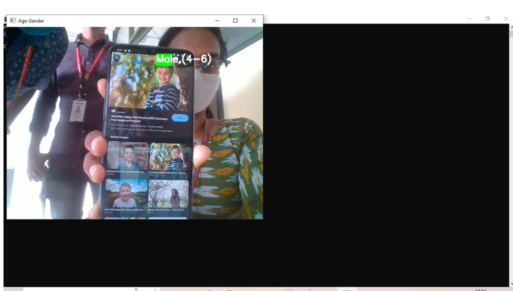

Fig. 8.

recently, this research has clearly focused on stunning images taken under laboratory conditions. For providing the correct and necessary settings for the environment of testing reflects how current real-world photography generally appears on social network platforms

We have compared our results with K. Venkateswara Rao et al.[1], here input image is taken from dataset for age estimation and gender prediction using CNN algorithm it resulted an accuracy of 73 percentage. In our paper we have developed age and gender recognition model using CNN that takes a live image of person as input through webcam. In this model the accuracy is 88 percentage.l

Person in Fig. 5 is male with age 72, the model predicted gender to be male and age range as 60-100. Person in Fig. 6 is female with age 21, the model has shown gender as female and age range as 21-32. Age of the person Fig. 8 is 5 with gender male, the model predicted it as male and age range between 4-6. Therefore through this model we can see the age and gender recognition is done accurately where image is being take from web cam i.e., live image as input, and using CNN for classification

References

[1] Age Estimation and Gender prediction of Unfiltered Faces Using Deep Learning Techniques K. Venkateswara Rao1, G. Harika2, A. Blessy3, D. Vishnu Vardhan Reddy4 B. Rakesh5

[2] H. Han, C. Otto, and A. K. Jain, "Age estimation from face images: Human vs. machine per-formance," in Proc. Int. Conf. BTAS, Jun. 2013, pp. 1–8.

[3] UTKFace. (n.d.). Retrieved July 14, 2020, from http://aicip.eecs.utk.edu/wiki/UTKFace

[4] IMDB-WIKI – 500k+ face images with age and gender labels. (n.d.). Retrieved July 14, 2020, from https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/

[5] Cao, Q, Shen, L., Xie, W., Parkhi, O. M, Zisserman, A. (2018). VGGFace2 A Dataset for Recognising Faces across Pose and Age. 2018 13th IEEE International Conference on Automatic Face GestureRecognition (FG 2018). doi:10.1109/fg.2018.00020

[6] Akhand M. A. Sayim, M. I., Roy, S., Siddique, N.(2020).Human Age Prediction from Facial Image Using Transfer Learning in Deep Convolutional Neural Networks. Proceedings of International Joint Conference on Computational Intelligence Algorithms for Intelligent Systems, 217-229. doi:10.1007/978-981-15-3607-6-17

[7] Hu, L., Li, Z., amp; Liu, H. (2015).Age Group Estimation on Single Face Image Using Blocking ULBP and SVM. Proceedings of the 2015 Chinese Intelligent Automation Confer-ence Lecture Notes in Electrical Engineering, 431-438. doi:10.1007/978-3-662-46469-4-46

[8] A. Fariza, Mu'arifin and A. Z. Arifin, "Age Estimation System Using Deep Residual Network Classification Method," 2019 International Electronics Symposium (IES), Surabaya, Indonesia, 2019, pp. 607-611,doi: 10.1109/ELECSYM.2019.8901521.

[9] Rothe, Rasmus, Radu Timofte, and Luc Van Gool. "Dex: Deep expectation of apparent age from a single image." Proceedings of the IEEE International Conference on Computer Vision Workshops. 2015.

[10] Reddi, S. J., Kale, S., Kumar, S. (2018). On the convergence of Adam and beyond. ICLR 2018.

[11] Viola, P., amp; Jones, M. (n.d.). Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision andPattern Recognition. CVPR 2001. doi:10.1109/cvpr.2001.990517

[12] A. Kumar and F. Shaik," Importance of Image Processing", 2016
[13] Chenjing Yan, Congyan Lang, Tao Wang, Xuetao Du, and Chen Zhang," Age Estimation Based on Convolutional Neural Network", 2014 Springer International Publishing Switzerland
[14] Aditya K. Saxena, Shweta Sharma and Vijay K. Chaurasiya," Neural Network based Human Age-group Estimation in Curvelet Domain", 2015 The Authors Published by Elsevier B.V.
[15] Dong Yi (B), Zhen Lei, and Stan Z. Li," Age Estimation by Multi-scale Convolutional Network", 2015 Springer International Publishing Switzerland
[16] A. Ultsch, G. Guimaraes, D. Korus, H. Li," Knowledge Extraction from Artificial Neural Networks and Applications", 1993 Springer
[17] Eran Eidinger, Roee Enbar, and Tal Hassner," Age and Gender Estimation of Unfiltered Faces", 2014 IEEE
[18] Yunjo Lee, Cheryl L. Grady, Claudine Habak, Hugh R. Wilson, and Morris Moscovitch," Face Processing Changes in Normal Aging Revealed by fMRI Adaptation", 2011 Massachusetts Institute of Technology
[19] R. Begg1 and J. Kamruzzaman," Neural networks for detection and classification of walking pattern changes due to ageing", 2006 ACPSEM/EA
[20] Hang Qi and Liqing Zhang," Age Classification System with ICA Based Local Facial Features", 2009 Springer-Verlag Berlin Heidelberg
[21] Kensuke Mitsukura, Yasue Mitsukura, Minoru Fukumi, and Sigeru Omatu," An Age Estimation System Using the Neural Network", 2003 Springer-Verlag Berlin Heidelberg
[22] Chao Yin and Xin Geng," Facial Age Estimation by Conditional Probability Neural Network", 2012 Springer-Verlag Berlin Heidelberg
[23] Sarah N. Kohail," Using Artificial Neural Network for Human Age Estimation Based on Facial Images", 2012 International Conference on Innovations in Information Technology
[24] Agbo-Ajala, O., Viriri, S. (2020). Deeply learned classifiers for age and gender predictions of unfiltered faces. The Scientific World Journal, 2020.
[25] Aruleba, I., Viriri, S. (2021, June). Deep Learning for Age Estimation Using EfficientNet.In International Work-Conference on Artificial Neural Networks (pp. 407-419). Springer, Cham
[26] Bekhouche, S. E., Ouafi, A., Benlamoudi, A., Taleb-Ahmed, A., Hadid, A. (2015, May). Facial age estimation and gender classification using multi-level local phase quantization. In 2015 3rd International Conference on Control, Engineering Information Technology (CEIT) (pp. 1-4). IEEE.
[27] Eidinger, E., Enbar, R., Hassner, T. (2014). Age and gender estimation of unfiltered faces. IEEE Transactions on information forensics and security, 9(12), 2170-2179.
[28] Iloanusi, O. N., Mbah, C. C., Ejiogu, U., Ezichi, S. I., Koburu, J., Ezika, I. J. (2019). Gender and age group classification from multiple soft biometrics traits. International Journal of Biometrics, 11(4), 409- 424.
[29] Hosseini, S., Lee, S. H., Kwon, H. J., Koo, H. I., Cho, N. I. (2018, January). Age and gender classification using wide convolutional neural network and Gabor filter. In 2018 International Workshop on Advanced Image Technology (IWAIT) (pp. 1-3). IEEE.
[30] Rafique, I., Hamid, A., Naseer, S., Asad, M., Awais, M., Yasir, T. (2019, November). Age and gender prediction using deep convolutional neural networks. In 2019 International conference on innovative computing (ICIC) (pp. 1-6). IEEE
[31] Kwon, Y. H., da Vitoria Lobo, N. (1999). Age classification from facial images. Computer vision and image understanding, 74(1), 1-21.
[32] Zhang, K., Gao, C., Guo, L., Sun, M., Yuan, X., Han, T. X., ... Li, B. (2017). Age groupand gender estimation in the wild with deep RoR architecture. IEEE Access, 5, 22492-22503.
[33] Liu, X., Li, S., Kan, M., Zhang, J., Wu, S., Liu, W., ... Chen, X. (2015). Agenet: Deeply learned regressor and classifier for robust apparent age estimation. In Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 16-24).
[34] Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.
[35] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
[36] Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and¡ 0.5 MB model size. arXiv preprint arXiv:1602.07360.