Age, Gender and Emotion Estimation Using Deep Learning

Chapter · July 2022

DOI: 10.1007/978-981-16-9113-3_6

Mala Saraswat
Bennett University

Age, Gender and Emotion Estimation Using Deep Learning

Mala Saraswat, Praveen Gupta, Ravi Prakash Yadav, Rahul Yadav, and Sahil Sonkar

Abstract Age, gender and emotion estimation plays very important role in intelligent applications such as human–computer interaction, access control, healthcare and marketing intelligence. To make computer demonstrate about people age, gender and emotion, lot of research has been conducted. However, it is yet a long way behind the human vision framework. This paper proposes and build an automatic age, gender and emotion estimation towards human faces. This estimation plays a significant part in computer vision and pattern recognition. Non-verbal specialized techniques like facial appearances, eye variation and gestures are utilized in numerous applications of human computer interconnections. This paper proposes a convolutional neural network (CNN)-based engineering architecture for age, gender and emotion classification. The model is trained to categorize input images into eight groups of age, two groups of gender and six groups will be used for the emotion. Basically, our approach shows better accuracy in age, gender and emotion classification compared with different classifier-based methods. In computer modeling the planning is to predict human emotions using CNN and observe changes occurred on emotional intensity. For extracting the features of images preprocessing algorithm that is known as Voila-Jones calculation is done. Experiments conducted using different data-sets: FER13 using our proposed approach provides accuracy of 81% for emotion estimation, for age 79% and gender accuracy is 75%.

Keywords Face detection · Viola-Jones · Face recognition · Deep learning · CNN

1 Introduction

Age, gender and facial expression are the main features to recognize a person. It also plays a very important role in social interconnections. Making age, gender and facial

M. Saraswat (✉) · P. Gupta · R. P. Yadav · R. Yadav · S. Sonkar
ABES Engineering College, Ghaziabad, Uttar Pradesh, India

M. Saraswat et al. (eds.), *Congress on Intelligent Systems*, Lecture Notes on Data Engineering and Communications Technologies 111,
https://doi.org/10.1007/978-981-16-9113-3_6

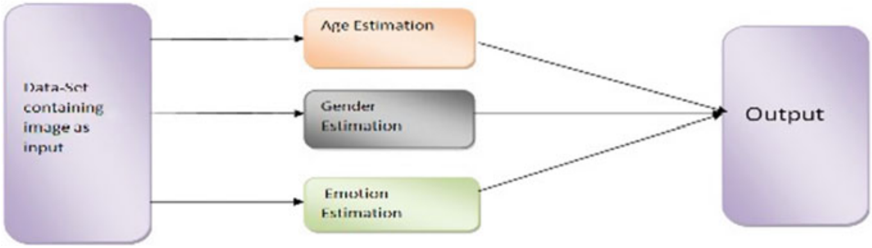

Fig. 1 Overall flow diagram of age, gender and emotion estimation

expression approximation from a single face-image is an important task in intelligent applications such as visual surveillance, law enforcement human–computer interaction, marketing intelligence and access control.

The outcome of the research is to develop a web application which will be acting as interface for users. It gives the facility to the user to upload an image of a face of a person and on the basis of that person age, gender and emotion will be estimated.

The process involved in our manuscript is shown in Fig. 1.

Section 2 discusses the related work, proposed methodology is described in Sect. 3, Sect. 4 contains results and experiments followed by Sect. 5 for conclusions and future work.

2 Related Work

Much research has been conducted for age, gender and emotion estimation. This section focuses on some related work.

A system for Facial Expression Recognition (FER) has been created using convolutional neural network (CNN). It plays a significant part in identifying facial expressions, which would then be able to enable neural networks having under 10 layers to complete with deeper networks for feeling acknowledgment introduced [1]. Face acknowledgment was accomplished effectively however they are influenced by light, facial expression, pose, face containing nose, face vocal points, eyebrows, mouth length using DLIB in OpenCV [2, 3]. In their model, during preprocessing Adaboost technique is used to eliminate unessential features and for extracting Haar like features as input to convolutional neural network (CNN) model for preprocessing Viola-Jones calculation is used [4, 5].

The deep model is prepared on an enormous data-set of four million pictures for the task of face acknowledgment. Above model fills in as the backbone to our facial property recognizer and is utilized to fine-tune networks for four tasks: age estimation, gender estimation and emotion estimation. Images are collected from different sources which are used in different tasks. There are 40,000 people images that are more than 4 million images are used for this model. Every picture is labeled in order to age, gender and data part is annotated with emotion. These pictures are subsequently managed utilizing a semi-robotized measure with a group of human annotators tuned in loop. The pictures are then pre-processed next to extract the faces and adjust them. The adjusted images are then fed to our proprietary deep network for training [6].

In beginning, the majority of the techniques in a very long time and gender assessment was carefully assembled and they are physically engaged designing of the facial highlights from the face. The first strategy for age assessment focusing on mathematical highlights of the face was created by Kwon and Lobo [7]. This strategy is utilized to decide the proportions among various elements of facial highlights [7].

A dynamic appearance model-based which included both the mathematical and surface highlights was proposed by Lanitis et al. [8].

From 2007, the greater part of the methodologies was utilizing physically planned highlights for the assessment task. These highlights consist of Spatially Flexible Patch, Gabor [9], local binary patterns and biological inspired features. By these features, age and gender of facial pictures are classified into two methods: Classification method. Regression's method.

In most recent years, there is an increase in researchers which use CNN for gender and age classification. CNN is utilized to group the gender and age of unfiltered face pictures [10]. The CNN methods are adopted for high-end computer machines and training on sufficiently large data for the classification task.

A model based on CNN was proposed by Levi et al. in 2015 in which there are five layers, which is partitioned into three convolutional and two completely associated layers to foresee the age of real-world face pictures [11]. Yi et al. [12] also, applied a start to finish perform multiple tasks CNN framework that learns a more profound design and the boundaries expected to address an age, a gender and an emotion characterization task in their paper. A model based on CNN was developed in 2018 by Liu et al. [13] that employed a multi-class focal loss function.

A hybrid CNN structure was presented in 2018 by Duan et al. [14] for age and sexual orientation arrangement which incorporates CNN and extreme learning machine (ELM) in which the CNN removes the highlights from the information pictures while ELM groups the transitional outcomes.

Computer used many approaches to detect, extract and recognize human facial features and expressions. Facial articulation location is attainable with low goal has appeared by Zhang. Zhang furthermore shows that there is important disposition information is encoded inside the internal facial features. These interior facial features allow visible presentation affirmation to be viably performed with commonly low computational essentials.

Khaorindish et al. in their paper proposed a hybrid model, with consideration of both CNN and SVM model to detect and classify the tumor detection and classifying MRI brain images [15]. Gupta et al. [16] used statistical techniques for identification of age, gender and face [17].

3 Proposed Methodology

The flowchart of proposed methodology is shown in Fig. 2. This section explains in detail the various steps of the proposed methodology in subsections.

3.1 Image or Data-Set Collection

Data-set used for gender, age and facial emotion detection are:

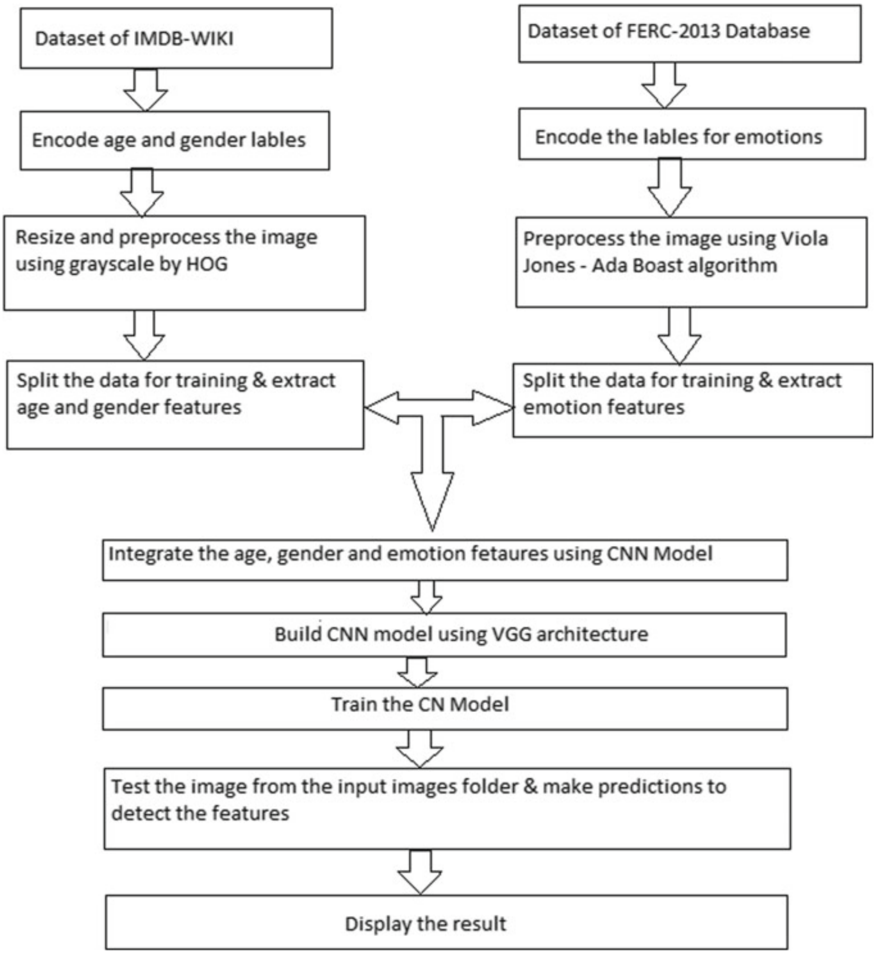

Fig. 2 Flowchart of proposed methodology

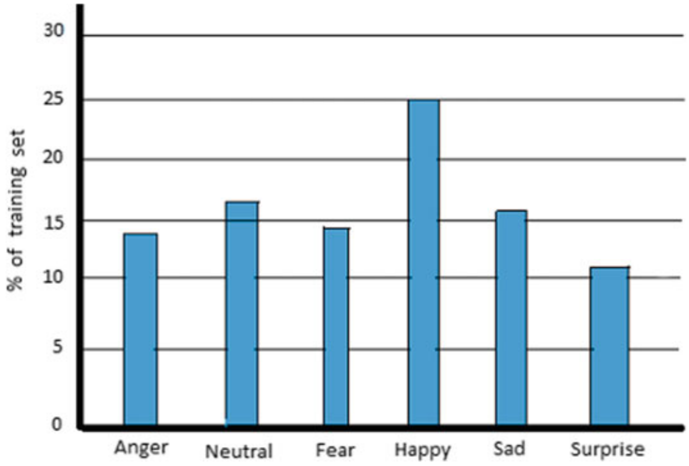

Fig. 3 Distribution of different emotions across the data-set

FER 2013: Facial expression recognition data-set used for emotion detection and classification. In this data-set, it evaluated many different preprocessing techniques for developing custom CNN model. Distribution of different emotions across the data-set are shown in Fig. 3.

IMDB-WIKI: This data-set used for age and gender detection. It is elaborated analysis of image for age and gender. Distribution of age and gender in IMDB-WIKI database are shown in Figs. 4 and 5, respectively.

3.2 Function to Encode the Labels

After collecting the data-set, we assign labels to this image-set using one-hot encode because CNN does not work with categorical data-variables that contain label values rather than numeric values. To overcome this problem of CNN, we use a one-hot encoding algorithm to convert these label values into numeric values which will be easily processed by CNN (Table 1).

3.3 Resizing and Preprocessing the Data

Resizing of all the images to fix pixel 256 $\times$ 256 values are done and preprocessing is done by converting them into grayscale using Histogram of Gradient (HOG) algorithm. For emotions, image preprocessing is done by using Viola-Jones–Ada Boost calculation to remove Haar like features uniquely utilized for identifying feelings.

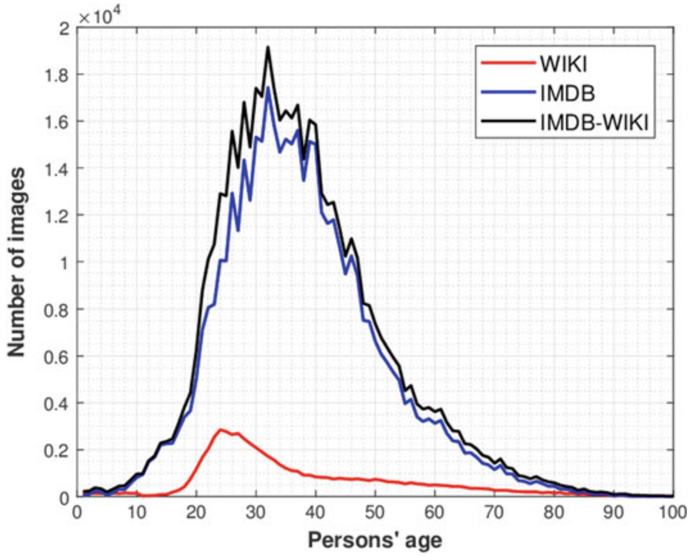

Fig. 4 Distribution of age in IMDB-WIKI database

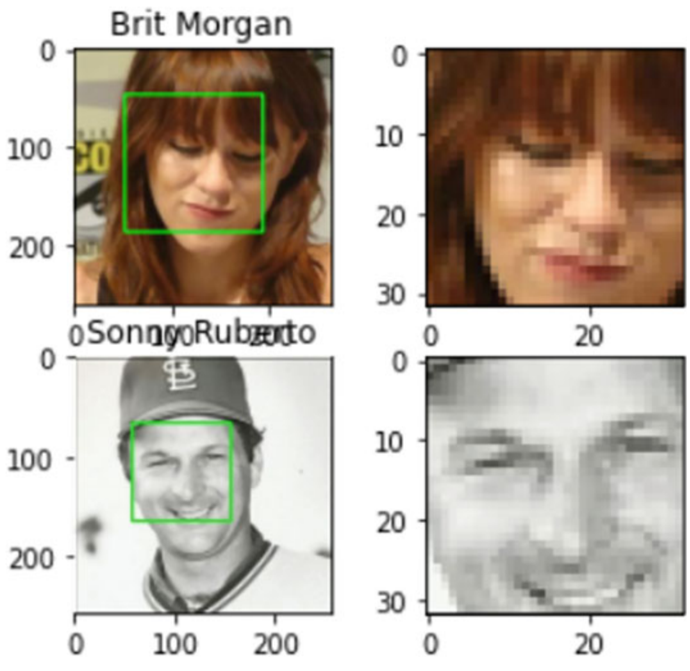

Fig. 5 Displaying a sample of extracted faces. *Source* IMDB-WIKI data-set

Table 1 Distribution of gender in IMDB-WIKI database

	WIKI	IMDB	IMDB-WIKI
Male	47,063	263,214	310,277
Female	12,622	189,047	201,669
Total	59,685	452,261	511,946

Viola-Jones adopts a gathering strategy. Which means is that Viola-Jones utilizes numerous elective classifiers, each gazing at an unmistakable bit of the picture. Each individual classifier is more fragile (less right, creates all the more bogus positives, and so on) than a definitive classifier because of its taking in less data. The image is reshaped in such a way that it only considers the facial features of the image as shown in Fig. 5.

3.4 Extracting and Integrating the Features

After preprocessing, the facial features such as eyebrows and distance between them, nose, mouth length and face landmark points are extracted using the DLIB library which is present in OpenCV. Then age, gender and emotion features are as integrated as one for training.

3.5 Training and Testing

The CNN model is built by using VGG-16 architecture. The CNN model is then trained using epochs, where each epoch contains a certain number of training images. To remove distorted and unwanted images, the loss Gauss function is used. For testing, the input image is given by the user. The model makes the predictions to estimate age, gender and emotion of that input image by comparing with the trained images.

3.6 Output

In the output stage, we apply a similar feature extraction process to the new pictures and we pass the features to the trained artificial intelligence calculation to anticipate the mark.

4 Experiment and Results

Our CNN model is implemented using Caffe open-source framework [18]. We implemented our model on different person in Figs. 6 and 7. In Fig. 6, a female is miscalculated as male and in Fig. 7, a female which has to be estimated in her age around 50–60 is miscalculated as (21–35).

Fig. 6 Miscalculation of gender of female

Fig. 7 Miscalculation of age of female

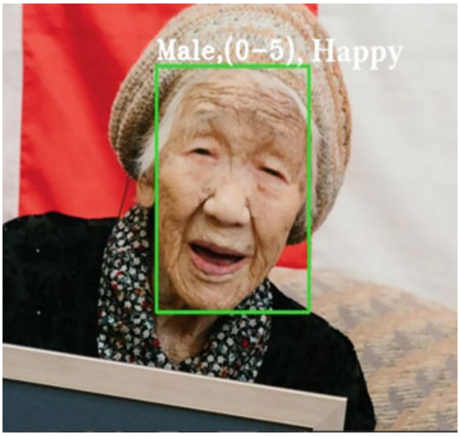

Fig. 8 Miscalculated gender and age of female

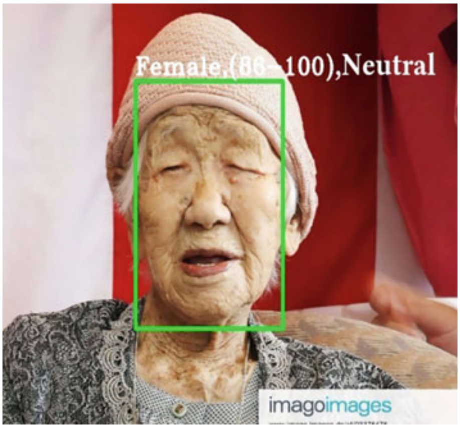

Fig. 9 Miscalculated emotion but gender and age is correct for the same female

We have implemented our model on same person in Figs. 8 and 9. In the Fig. 8, gender and age are miscalculated for the female in her age (60–100) as male and age (0–5), respectively and in the Fig. 9, for the same female emotion is miscalculated but gender and age are correct as female and age (60–100), respectively.

We provide few examples of age, gender and emotion miscalculation in Figs. 6, 7, 9 and 10. These shows that many of the mistakes made by our CNN model. Gender estimation mistakes are frequently occurred for pictures of infants or exceptionally little youngsters and in old individual where obvious gender attributes are not yet visible. Emotion estimation for pictures of infants or exceptionally little youngsters are also miscalculated obviously since appropriate emotion attributes are not visible in the image.

To benchmark this investigation utilizing best in class age assessment methods, Table 2 gives a comparison of accuracy accomplished through this current work's

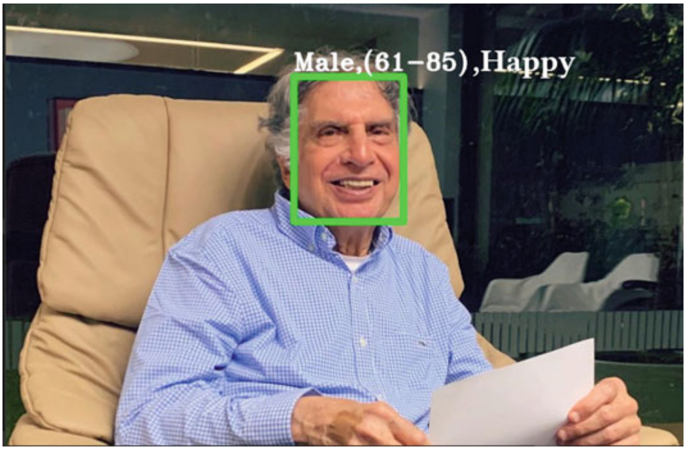

Fig. 10 An image of Mr. Ratan Tata in which age, gender and emotion is estimated

Table 2 Proposed estimation scheme versus other method of feature selection

Research papers	Feature extraction	Age/gender/emotion estimation	Data-set	Accuracy
Han and Jain [19]	Support Vector Machine (SVM)	Age	FG-NET	87.15
Hasan and Mahdi [20]	Local Binar Patterns (LBP)	Age	FG-NET	93.81
Sinha [21]	Principal Component Analysis (PCA)	Gender	IIT Kanpur	86.6 and 93.3
Singh and Shokeen [22]	LBP and HOG	Gender	Indian face	89.4 and 95.56
Sharma and Dutta [23]	HOG + Voila-Jones Algorithm	Age, Gender and Emotion	FER13	72.53. 98.9 and 70
Ramesh and Venugopal [24]	FEBFRAC Algorithm	Age, Gender	NA	90 and 95
Pao [25]	HOG	Emotion	NA	81
Hwan and Hoon [26]	Fuzzy classifier	Emotion	NA	74
Dehghan [27]	Sight bound	Age, Gender and Emotion	IMDb-Wiki and	96.2, 90.3 and 97.1
Proposed Method	HOG + Viola-Jones + Adaboost Algorithm	Age, Gender and Emotion	IMDb-Wiki and FER13	79, 75 and 81

proposed age, gender and emotion estimation scheme and other recently-published works in the field of age, gender and emotion estimation.

A snapshot of our proposed method CNN model is shown in Fig. 10, showing estimated gender, an age and an emotion.

We have prepared the framework on certain VIPs and a portion of our companions, we noticed that sometimes that the two face vectors of two distinct individuals were even found nearer then one another (Euclidean distance) than two photographs of one individual. From what we tried, utilizing 5+ photographs per individual with clean, clear, forward-looking faces would create better outcomes for character estimates.

5 Conclusions and Future Work

By our proposed methodology a complete overview of the state-of-the-art technique for gender, age and emotion classification has been inspected and talked about through face pictures. In this research paper, various data-set and various algorithms have been proposed by their researchers. For the proposed method accuracy for emotion HOG + Viola-Jones algorithm is 81% in FER 2013 data-set and for accuracy of age and gender estimation are 79% and 75%, respectively.

In this paper, we have utilized pre-built Caffe models for age and gender prediction in OpenCV in Python. This one is based on regular VGG, although its keras execution is based on VGG-Face. We have likewise developed CNN model to perceive looks of people. Handling recognized faces rather than the whole picture would build accuracy. Application of our methodology in different areas like security, lie detection, automated tutoring system, forensic, face detection systems, etc.

References

1. Minaee S, Abdolrashidi A (2019) Deep-emotion: facial expression recognition using attentional convolutional network. In: 2019 3rd IEEE International conference on cybernetics (CYBCONF). https://doi.org/10.1109/cybconf.2017.7985780
2. Salihbasi A, Orehovacki T (2019) Development of android application for gender, age and face recognition using opencv. MIPRO 2019, ISSN: 1057-7149
3. Xie J-C, Pun C-M (2019) Age and gender classification using convolutional neural networks. In: IEEE workshop on analysis and modelling of faces & gestures', IEEE transactions on information forensics and security. ISSN: 1556-6013
4. Levi G, Hassner T (2017) Facial emotion analysis using deep convolutional neural network. In: International conference on signal processing and communication (ICSPC), Coimbatore
5. Mohammadpour M, Mohammed S, Hashemi R, Khaliliardali H, Alyan Nezhadi MA (2017) Facial emotion recognition using deep convolutional networks. In: 2017 4th International conference on knowledge-based engineering and innovation (KBEI), 22 December 2017
6. Dehghan A, Ortiz EG, Shu G, Zain Masood S (2017) DAGER: Deep age, gender and emotion recognition using convolutional neural network
7. Kwon YH, Da Vitoria Lobo N (1999) Age classification from facial images. Comput Vis Image Underst 74(1):1–21

8. Lanitis A, Draganova C, Christodoulou C (2004) Comparing different classifiers for automatic age estimation. IEEE Trans Syst Man Cyber Part B (Cybernetics) 34(1):621–628
9. Gao F, Ai H (2009) Face age classification on consumer images with gabor feature and fuzzy LDA method. In: Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and Lecture notes in bioinformatics), vol 5558, pp 132–141. LNCS, Springer, Berlin, Germany
10. Szegedy C, Liu W, Jia Y et al (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Boston, MA, USA, pp 1–9
11. Levi G, Hassncer T (2015) Age and gender classification using convolutional neural networks. In: Proceedings of the 2015 IEEE Conference on computer vision and pattern recognition workshops (CVPRW). Boston, MA, USA, pp 34–42
12. Yi D, Lei Z, Li SZ (2015) Age estimation by multi-scale convolutional network, Computer vision–ACCV 2014. Springer, Berlin, Germany
13. Liu W, Chen L, Chen Y (2018) Age classification using convolutional neural networks with the multi-class focal loss. IOP Conf. Ser.: Mater. Sci. Eng. 428(1)
14. Duan M, Li K, Yang C, Li K (2018) A hybrid deep learning CNN–ELM for age and gender classification. Neurocomputing 275:448–461
15. Khairandish MO et al (2021) A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images. IRBM
16. Gupta R, Kumar S, Yadav P, Shrivastava S (2018) Identification of age, gender, & race SMT (scare, marks, tattoos) from unconstrained facial images using statistical techniques. In: 2018 International conference on smart computing and electronic enterprise (ICSCEE). IEEE, pp 1–8
17. Gupta R, Yadav P, Kumar S (2017) Race identification from facial images using statistical techniques. J Stat Manag Syst 20(4):723–730
18. Zhang Z et al (2007) Joint face and head tracking inside multi-camera smart rooms. Signal Image Video Process 1(2):163–178
19. Han H, Jain AK (2014) Age, gender and race estimation from unconstrained face images. MSU Technical Report (MSU-CSE-14-5), 87, 27. Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, USA
20. Hasan NF, Mahdi SQ (2020) Facial feature extraction for age estimation based on a neural network. Department of Computer Engineering Technology, Middle Technical University, Baghdad, Iraq
21. Sinha C (2013) Gender classification from facial images using PCA and SVM. Department of Biotechnology and Medical Engineering, National Institute of Technology Rourkela, Rourkela, 769008, India
22. Singh V, Shokeen V (2013) Comparison of feature extraction algorithms for gender classification from face images. ECE Department Amity University Noida, India
23. Sharma DJ, Dutta S (2020) Real-time emotion, gender, age detection using CNN. Student, School of Computing Sciences(IT), The Assam Kaziranga University, Jorhat, Assam, India
24. Ramesha K, Raja KB, Venugopal KR. Feature extraction based face recognition, gender and age classification. Department of Telecommunication Engineering, Vemana Institute of Technology, Bangalore, 560034, India
25. Poa J. Emotion detection through facial feature recognition. Stanford University
26. Hwan M, Hoon Y. Emotion detection algorithm using frontal face image. Department of Electrical and Electronic Engineering, Yonsei University Korea
27. Dehghan A, Ortiz EG. Deep age, gender and emotion recognition using convolutional neural networks. Computer Vision Lab, Sighthound Inc., Winter Park