2023 International Conference on Computer Communication and Informatics (ICCCI), Jan 23-25, 2023, Coimbatore, India

Age and Gender Prediction using Adaptive Gamma Correction and Convolutional Neural Network

Anweasha Saha
Dept. of Computing Technologies
SRM University, Kattankulathur
Chennai, India
as6009@srmist.edu.in

Nithish Kumar S
Dept. of Computing Technologies
SRM University, Kattankulathur
Chennai, India
sk7649@srmist.edu.in

Nithyakani P
Dept. of Computing Technologies
SRM University, Kattankulathur
Chennai, India
nithyakp@srmist.edu.in

***Abstract*— The classification of age and gender has drawn increased attention recently because of its significance in creating user-friendly intelligent systems. In the domains of image processing and computer vision, determining age from a single facial image has proven a challenging job. Convolutional Neural Network (CNN) based techniques have been frequently adopted for the classification problem in the recent past because of their precise results in facial analysis. This study incorporates an end-to-end CNN approach with the addition of a key pre-processing step for image contrast enhancement, which was done via an adaptive gamma correction technique to produce precise gender and age group classification of real-world faces. The complete feature extraction and classification processes are included in the two-level CNN architecture. The feature extraction task pulls features that are related to gender and age while the classification assigns the facial photographs to the proper gender and age group. The proposed network has been trained and tested on the Adience (original) dataset. The results of the experiment seem to back up the claim that the proposed model is better at classifying people by gender and age when the Adience benchmark for classification is used.**

Keywords—Age Identification, Facial Images, Gender Classification, Convolution Neural Network, Deep Learning

I. INTRODUCTION

A variety of attributes, obtained from facial appearances, play a significant role in social interaction with others. Among these, age and gender are the most fundamental features to estimate others' emotions or characteristics for proper and efficient communication. Ageing is an inevitable process of natural change that adds to human looks. As such, automatic age and gender detection through facial analysis have grown to be an important research topic over the past years. This can be very beneficial for various real-world applications, starting from security systems like biometrics and video surveillance, to interactive systems and medical diagnosis. Although researchers face many challenges in building an accurate and robust system for age and gender classification. As several factors influence the ageing process and it's never the same for any two persons, it is hard even for a human to predict someone's exact age. The same applies to machine learning models.

Due to its potent capacity to evaluate and extract attributes to improve the accuracy of image classification, convolutional neural networks focused on deep learning have recently made significant advances in a variety of fields. Due to the network's high accuracy and pattern recognition capability, most organizations are interested in using this algorithm for their own profit. Additionally, to make it more credible, there are now technologies for explaining AI's solution. In order to illustrate all the preliminary actions taken before implementing the model, the basic framework of the entire architecture in this paper is outlined. Pre-processing steps such as image resizing, image enhancement, image segmentation, and, most importantly, face detection and face alignment are all included in these modules. The main contribution involves the proposal of a two-level CNN model for categorizing age and gender-related data with adaptive gamma correction for facial image enhancement, which is followed up with a presentation of the model's results and the final conclusion.

II. RELATED WORK

Age prediction has proved to be a largely demanding affair in the field of facial analysis. In the past, age estimation was utilised for manually extracting facial characteristics. However, more recently, CNN methods are being favoured because of the successful training of CNN directly on age datasets.

Transfer learning is implemented using two pre-trained models VGG19 and VGGFace with several training techniques [1]. Using distinct male and female age models, a hierarchy of deep CNNs has been explored that first groups subjects by gender before estimating the age. On the MORPH-II dataset, the proposed technique has obtained a 98.7% accuracy for gender classification and a mean absolute error of 4.1 years for age prediction. For simultaneous age and gender recognition, J.H. Lee et al., has developed a lightweight multi-task CNN (LMTCNN) that is more effective than baseline multi-task CNN approaches [2]. In order to shorten the inference time and reduce the size of the model, LMTCNN employs depth-wise separable convolution. On the Adience dataset, the obtained gender and age classification accuracies are 85.16% and 44.26%, respectively. The same year, A. Das et al., proposed the usage of a Multi-Task Convolution Neural Network (MTCNN) that employs joint dynamic loss weight adjustment to determine age and gender [3]. For the UTKFace dataset, the mean gender classification accuracy is 98.23%, while for the BEFA challenge dataset, it is 93.72%. On the BEFA challenge dataset, the age estimation accuracy is 71.83%, while on the UTKFace dataset, it is 70.1%. The method put forward by S.

Hosseini et al., uses back-propagation in a CNN-based architecture to employ Gabor filter responses as input for combined age and gender identification [4]. The model produced accuracy levels of 61.3% for age recognition and 88.9% for gender classification on the Adience dataset.

A unique approach for fine-grained age estimation has been put out by K. Zhang et al., which is based on attention long short-term memory (AL) network [5]. The method creates AL-RoR networks or AL-ResNets by combining the residual network of residual network (RoR) models or residual networks (ResNets) with LSTM units. These networks considerably increase the accuracy by extracting the local characteristics from age-sensitive areas. For age group classification, an accuracy of 67.83% and 2.63 MAE are obtained on the Adience dataset and MORPH Album 2 dataset respectively. Insha Rafique et al., has implemented a simple 3-convolutional layer CNN model [6]. To bring about a noticeable improvement, HAAR Feature-based Cascade Classifiers have been used for training the model. The method yielded age and gender identification accuracy rates of 79.3% and 50.7%, respectively. The same year, Kyoungson Jhang et al., reported that in camera-based testing, a CNN named Googlenet that was trained using grayscale images exhibited noticeably better age group as well as gender prediction accuracy [7]. Apparently, grayscale image files appear to make CNN more resistant to alterations in lighting than RGB ones. For the Adience dataset, the gender classification accuracy is 78.10% and the exact accuracy for age estimation is 38.50%. A. V. Savchenko et al., suggested a two-stage method using a modified MobileNet, which was implemented in an Android application. Initially, the CNN network identifies age and gender while concurrently extracting facial features that are suitable for face recognition [8]. The extracted faces are then grouped in the second stage by hierarchical agglomerative clustering methods. On the UTKFace dataset, this technique achieves an accuracy of 94.1% for gender recognition and a mean absolute error (MAE) of 5.44 for age prediction.

Cao Hong Nga et al., has implemented a transfer learning technique to reduce the training time and increase the final accuracy [9]. They have used ImageNet pre-trained models by dividing them into different stages with decaying learning rate in schedule. With this approach, they have produced an accuracy of 91.09%. Mohammed Kamel Benkaddour et al., has presented three CNN models with different architecture depths, namely, CNN 1, CNN 2, and CNN 3 along with a varying number of filters, pooling, and convolutional layers [10]. Concerning gender classification, the best accuracy rates for both IMDb and WIKI datasets have been given by CNN 3, 94.49% and 93.56% respectively. To verify the results of this model, it has been implemented for class-wise age estimation, i.e., batches of 20-39 years, 40-59 years, and more than 60 years. A mean accuracy of 86.20% has been obtained on the IMDb dataset and 83.97% on the WIKI dataset. Earnest Paul Ijjina et al., in the same year explored the effectiveness of Wide Residual Networks on a dataset of surveillance video acquired from a garment store's CCTV cameras [11]. Their model is known as the Wide ResNet 16-8, which refers to the WRN variant with 16 layers of convolution and an 8-fold widening factor. The gender classification task is accurate at 82.926%, whereas the age prediction task is accurate at a percentage of 70.804. An innovative end-to-end CNN technique has been described in the work of O. Agbo-Ajala et al., to determine gender and age from unfiltered real-world faces [12]. The OIU-Adience dataset is used to assess the two-level CNN architecture. The proposed method achieved an exact accuracy of 83.1% for age estimation and 96.2% for gender recognition. In the approach suggested by Anirudh Ghildiyal et al., a CNN network is employed which has 3 convolutional layers, 2 completely interconnected layers, and a final output layer [13]. It manages to pull off an accuracy of up to 0.6170588.

Jayaprada S et al., has compared and tested two approaches, namely back propagation neural network and support vector machine [14]. The simpler model SVM performed better than BPNN (78% accuracy) on the Adience dataset with an accuracy of 90%. N. Shanthi et al., has proposed a deep convolutional neural network with just three fully connected layers and two associated layers each with the least number of neurons having the intention to reduce the possibility of overfitting [15]. With this approach, the model achieved a 96.2% accuracy for gender prediction and 57.51%accuracy for age estimation.

III. DATASETS USED

A. Adience Dataset

The Adience dataset, containing 17,856 photos across 2,284 subjects, is as close as it can get to real-world face imaging conditions. It comes with a binary gender label in addition to eight distinct age classes, split into five sections. The images were crawled from Flickr albums which were uploaded from smartphones without any filtering. As a result, the dataset has a diverse collection of faces having variations in all aspects such as pose, appearance, lighting, image quality and so on.

Adience dataset was used to train our network for the categorization of gender and age groups.

IV. PROPOSED WORK

The deeply learned classifiers which we have proposed in this work to identify age and gender group from real-world unfiltered facial images are described in this section below.

Before the facial pictures are supplied to the suggested network, our technique involves an image preparation step (face detection, landmark detection, and face alignment). Hence, the three main components of our system are image pre-processing, feature learning, and classification itself.

A. Image Processing

There is a huge necessity to pre-process the unfiltered real-world images so that our proposed network can tackle the classification task successfully. The majority of those facial images are non-frontal, nonaligned, and have different degrees of lighting, stance and surrounding conditions. As such, in order to get better performance, the raw images need to be detected and aligned and then fed into the model.

B. Face Detection

Face detection is the initial step in image pre-processing. We have made use of Haar Cascades, an open-source face detector. All input photos are rotated between -90° and 90° angles with a 5° step for the sake of detecting faces. Next, the detector will select the face image that produces the finest possible output. If the corresponding facial image is still not found after all of the changes have been made, the initial image will be scaled up, and the face detection process will be repeated as many times as necessary until it does locate a face. The upscaling aids in the identification of faces across all input photos.

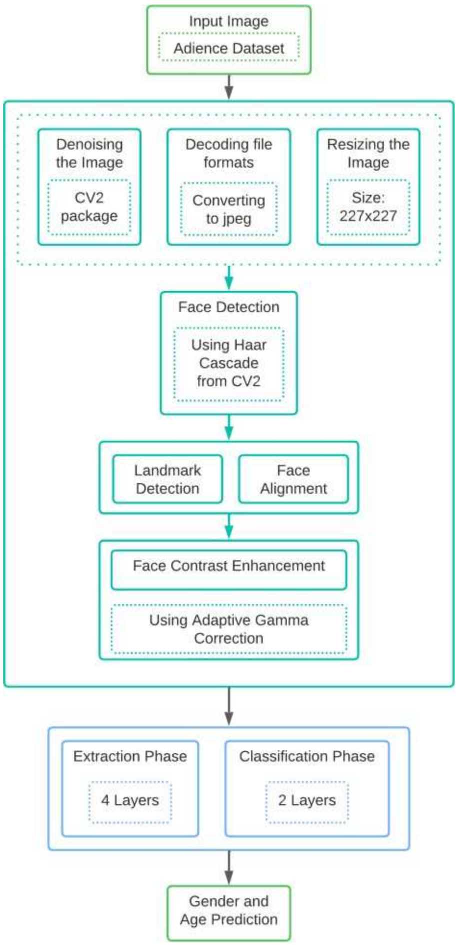

Fig. 1. Architecture of proposed pipeline

C. Landmark Detection

Next is facial landmark detection which predicts and tracks key points representing regions or landmarks on a human's face. This algorithm uses five landmark detection models, each of which has been trained in order to perform on one of the associated facial postures. These models are, namely, a frontal model, two full-profile models, and two half-profile models.

D. Face Alignment

The phase of the alignment of faces includes applying each of the five models to the detected faces. The model having the greatest confidence score is then transformed using an affine function to the predetermined ideal settings for those landmarks.

E. Image Contrast Enhancement using Adaptive Gamma Correction

Gamma correction is basically a non-linear approach, defined by a power law transform. As opposed to conventional linear techniques such as addition, subtraction and multiplication, gamma correction performs non-linear methods on each pixel with the intention of upgrading the saturation of the input image. Adaptive gamma correction (AGC), as proposed by Shanto Rahman et al., enhances the image contrast by dynamically adjusting the parameters of AGC based on the data extracted from the image [17].

F. CNN Framework

The CNN architecture proposed in this article is a six-layer network that consists of two fully connected layers and four convolutional layers. A sequential, end-to-end deep learning architecture, the CNN design comprises two steps, namely, feature extraction and classification. The four convolutional layers in the feature extraction phase have associated parameters, such as the number of filters, the stride, and the size of each filter's kernel. Convolutional layer, activation layer (rectified linear unit (ReLU)), batch normalisation, and a max-pooling layer. Whereas, the classification step of the model comprises two completely connected layers which are used to tackle the phase of classification.

The first fully connected layer consists of 512 neurons and is trailed with a ReLU, batch normalisation, and a dropout layer of 0.5. For classification tasks, the last and second fully connected layers generate 512 features that, depending on the task, are dense with 2 or 4 neurons.

Layer	Output Shape	Param #
sequential	(227, 227, 3)	0
Convolution1	(56, 56, 96)	14208
Batch_Normalization	(56, 56, 96)	384
Max pooling	(27, 27, 96)	0
Convolution 2	(27, 27, 256)	614656
batch_normalization_1	(27, 27, 256)	1024
Max_pooling_1	(13, 13, 256)	0
Convolution 3	(13, 13, 384)	885120
batch_normalization_2	(13, 13, 384)	1536
Max_pooling_2	(6, 6, 256)	0
Convolution _4	(6, 6, 256)	884992
batch_normalization_3	(6, 6, 256)	1024
Max_pooling_3	(2, 2, 256)	0
flatten	(1024)	0
dense	(512)	524800
dropout	(512)	0
dense_1	(512)	262656
dropout_1	(512)	0
dense_2	(1)	513

TABLE 1: MODEL SUMMARY

Specifically for the first dense layer of the age classification task, we have used an L2 regularisation factor of 0.01 to decay the weights and avoid the process from falling into local minima.

Both tasks are executed with different learning rates to balance the compatibility between the model complexity and the size of the dataset used. The values for the same are 0.01 and 0.001 for gender and age, respectively. Since the gender task is a binary classification, starting with a momentum value of 0.9 was required to avoid learning too much from the training data.

Finally, for both tasks, the stochastic gradient descent (SGD) algorithm is used as an optimizer, which has more popularity and is faster than any other optimizer.

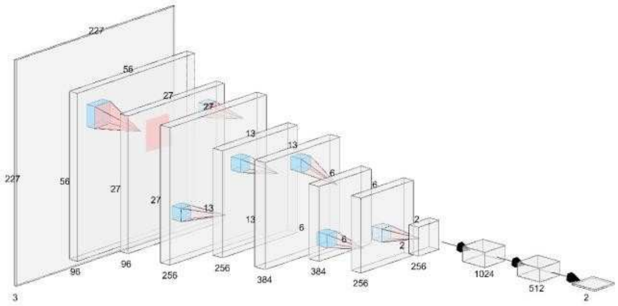

Fig. 2. Proposed CNN Architecture Diagram

The probability for each gender and age range is calculated using a cross-entropy, loss function and a softmax since the gender and age classification task is thought of as an end-to-end deep classification problem. The design of a few parameters of the CNN architecture has been defined below.

The probabilities for every class label are provided by the softmax classifier. The function for the same is illustrated in equation (1) below:

$$f_i(a) = \frac{e^{a_i}}{\sum_k e^{a_k}} \quad (1)$$

where f_i is the i-th element of the class score vector f, which will accept any vector of true scores in a.

For training binary and multi-class classifications for gender and age classifiers, a cross-entropy loss is required.

Equation (2) below is the cross-entropy function for binary classification:

$$(q) = -\frac{1}{N}\sum_{i=1}^{N}[a_i \log(p(a_i)) + (1 - a_i)\log(1 - pa_i)] \quad (2)$$

where a is the binary class label, 1 for the exact class and 0 otherwise, and p(a) is the predicted probability of the point being green over N points.

The representation of cross-entropy for multi-class classification is given as follows in equation (3):

$$H_{a^j}(a) = -\sum_j a'\log(a_j) \quad (3)$$

a_j denotes the probability of class j predicted value for class, whereas a_j' is the actual class probability.

V. RESULTS AND DISCUSSION

In this section, the findings that were obtained by testing the suggested approach on the Adience dataset are presented. The numerical results, that is, the accuracies for both the gender and age classification tasks have been mentioned below in Table 2. We have employed binary and multi-class classification for gender and age estimation respectively. Our CNN model achieved an accuracy of 84.94% for gender identification. The age prediction is an estimation that places the subject in the appropriate age range among the four categories, which are as follows: kid (0- 12), youth (13- 30), adult (31- 55), and old (56 and above). The model obtained an accuracy of 75.10% for this metric on the Adience dataset.

Previous Works		**Accuracy**	**Precision**	**Recall**
Insha Rafique et al., [6]	gender	50.7%	-	-
	age	79.3%	-	-
Kyoungson Jhang et al., [7]	gender	78.10%	-	-
	age	38.50%	-	-
Earnest Paul Ijjina et al., [11]	gender	82.92%	-	-
	age	70.80%	-	-
Proposed Work	**gender**	**84.94%**	**82.68%**	**84.50%**
	age	75.10%	63.65%	69.96%

TABLE 2: THE ACCURACY OF THE AGE AND GENDER CLASSIFICATION AS ACHIEVED BY OUR PROPOSED MODEL

Fig. 3. Confusion Matrix for Age Group Classification

The model obtained better results when the face alignment was proper, that is when the subject was facing the camera. Minor changes in the alignment degraded the performance of the model in some instances. The CNN architecture faced difficulties in detecting the side facial profile in those cases, thus affecting the numbers.

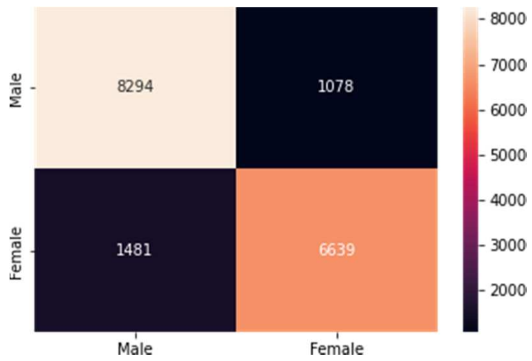

Fig. 4. Confusion Matrix for Gender Classification

Another reason for the lower age estimation accuracy might be that our model needs more data or larger datasets to train. Feeding more data into the model would improve the results.

VI. CONCLUSION AND FUTURE WORK

In this work, the problem of age and gender classification of real-life unfiltered facial images is addressed. The gender identification and age estimation tasks were posed as a binary and a multi-class classification problem respectively. A six-layer CNN architecture has been proposed and implemented for the same. Our proposed model is originally trained on the Adience dataset. It has achieved a gender accuracy of 84.94% and an accuracy of 75.10% for the age metrics. Haar Cascades has been employed to improve the results in the aspect of face detection.

For future works, we aim to increase the accuracy by making use of a larger dataset for fine-tuning our model. In addition, the model can be trained further because the loss functions are still decreasing. We wish to deploy this project as a real-time application in the near future.

REFERENCES

[1] P. Smith and C. Chen, "*Transfer learning with deep CNNs for gender recognition and age estimation*", *Proc. IEEE Int. Conf. Big Data (Big Data)*, pp. 2564-2571, Dec. 2018.

[2] J. H. Lee, Y. M. Chan, T. Y. Chen and C. S. Chen, "*Joint estimation of age and gender from unconstrained face images using lightweight multi-task CNN for mobile applications", Proc. IEEE Conf. Multimedia Inf. Process. Retr. (MIPR)*, Apr. 2018.

[3] A. Das, A. Dantcheva and F. Bremond, "*Mitigating bias in gender age and ethnicity classification: A multi-task convolution neural network approach*", *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*, pp. 1-13, 2018.

[4] S. Hosseini, S. H. Lee, H. J. Kwon, H. I. Koo and N. I. Cho, "*Age and gender classification using wide convolutional neural network and Gabor filter*", *Proc. Int. Workshop Adv. Image Technology (IWAIT)*, pp. 1-3, 2018.

[5] K. Zhang, N. Liu, X. Yuan, X. Guo, C. Gao, Z. Zhao, et al., "*Fine-grained age estimation in the wild with attention LSTM networks*", *IEEE Trans. Circuits Syst. Video Technology*, vol. 30, no. 9, pp. 3140-3152, Sep. 2020.

[6] Insha Rafique, Awais Hamid, Sheraz Naseer, Muhammad Asad, Muhammad Awais, Talha Yasir, "*Age and Gender Prediction using Deep Convolutional Neural Networks*", *2019 International Conference on Innovative Computing (ICIC)*, 2019

[7] Kyoungson Jhang, Junsoo Cho, "*CNN Training for Face Photo based Gender and Age Group Prediction with Camera*", *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2019

[8] A. V. Savchenko, "*Efficient facial representations for age gender and identity recognition in organizing photo albums using multi-output ConvNet*", *PeerJ Comput. Sci.*, vol. 5, pp. e197, Jun. 2019.

[9] Cao Hong Nga, Khai-Thinh Nguyen, Nghi C. Tran, Jia-Ching Wang, "*Transfer Learning for Gender and Age Prediction*", *2020 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*, 2020

[10] Mohammed Kamel Benkaddour, Sara Lahlali, Maroua Trabelsi, "*Human Age And Gender Classification using Convolutional Neural Network*", *2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH)*, 2020

[11] Earnest Paul Ijjina, Goutham Kanahasabai, Aniruddha Srinivas Joshi, "*Deep Learning based approach to detect Customer Age, Gender and Expression in Surveillance Video*", *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2020

[12] O. Agbo-Ajala and S. Viriri, "*Deeply learned classifiers for age and gender predictions of unfiltered faces*", *Sci. World J.*, vol. 2020, pp. 1-12, Apr. 2020.

[13] Anirudh Ghildiyal, Sachin Sharma, Ishita Verma, Urvi Marhatta, "*Age and Gender Predictions using Artificial Intelligence Algorithm*", *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, 2020

[14] Jayaprada S. Hiremath, Shantakumar B. Patil, Premjyoti S. Patil, "*Human Age and Gender Prediction using Machine Learning Algorithm*", *2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC)*, 2021

[15] N. Shanthi;P Yuvasri, S Vaishnavi, P Vidhya, "*Gender and Age Detection using Deep Convolutional Neural Networks*", *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2022

[16] A. Garain, B. Ray, P. K. Singh, A. Ahmadian, N. Senu and R. Sarkar, *"GRA_Net: A Deep Learning Model for Classification of Age and Gender From Facial Images,"* in IEEE Access, vol. 9, pp. 85672-85689, 2021, doi: 10.1109/ACCESS.2021.3085971.

[17] Rahman, S., Rahman, M.M., Abdullah-Al-Wadud, M. et al. *"An adaptive gamma correction for image enhancement."* J Image Video Proc. 2016, 35 (2016). https://doi.org/10.1186/s13640-016-0138-1