A Modified CNN for Age and Gender Prediction

Danish Ali[1] , Francesco Epifania[2], Naeem Ahmed[3] , Bilal Khan[4] , Luca Marconi[5], Ricardo Matamoros[6]

[1,4] GPGC Haripur affiliated with University of Haripur, Near new judiciary, Haripur, 22620, Pakistan
[2,6] Social Things srl, Italy
[3] UET, Taxila, Taxila, Pakistan
[5] University of Milano-Bicocca, Milan, Italy

Abstract

Age and gender information are essential for many real-world applications, such as social intelligence, biometric identity verification, video surveillance, human-computer interaction, digital consumer, crowd behavior analysis, online marketing, item recommendation, and many more. This study intends to employ deep learning technology in the prediction process, effective accuracy, and predictive mining and assess it in order to obtain the best outcomes of prediction and get around the issues of time, accuracy, and processing load. In this multi-task learning problem, age and gender are predicted concurrently with the help of a single Convolutional neural network with two heads (output branches). The model has 95% accuracy for gender classifier and 92% accuracy for age classifier. The pro-posed model uses the computing resources (RAM, CPU, and GPU) in a much more optimized manner and the computing cost is also lower.

Keywords

Convolutional Neural Network, Recognition System, Gender Prediction, Age Prediction

1. Introduction

For many real-world applications, including social intelligence, biometric identity verification, video surveillance, human-computer interface, digital consumer, crowd behavior analysis, online marketing, item suggestion, and many more, age and gen-der data are crucial [1, 14]. No matter how widespread their uses, being able to automatically determine age and gender from face pictures is a very difficult problem [15]. This is especially true given the various sources of intra-class variations at peo-ple's facial images, which restricts the use of these models in real-world programs [2, 16]. In the past several years, a lot of works have been offered for predicting age and gender [3]. Recent research has focused in particular on using a classifier after manually extracting face information from photos [17]. Nonetheless, because to the out-standing success of deep learning models in several computer vision issues over the past few years [18], the majority of the more recent efforts on age and gender predictions have turned toward models based on deep neural networks [4, 19].

As aim to propose a deep learning system in this study to jointly estimate the age and gender from facial images. Given the intuition that a few neighboring parts of the face provide very obvious messages regarding a person's age and gender [20] (inclusive of beard and moustache for male, and wrinkles around eyes and mouth for age) [5]. Employ a single version using a multi-task learning approach to collectively estimate both gender and age bucket since estimating age and gender from faces is highly correlated [6]. Additionally, as knowing a person's gender helps us estimate their age more accurately, add the predicted gender output to the age-prediction branch's feature [21].

In order to accurately anticipate the future and learn more about a specific man or woman, studies in the biometric field, including human face recognition applications, focus on gender and age

[1]2nd Italian Workshop on Artificial Intelligence and Applications for Business and Industries - AIABI | co-located with AI*IA 2022
EMAIL: danishalikhan545@gmail.com (A. 1); francesco.epifania@socialthingum.com (A. 2); Naeem.uoh@gmail.com (A. 3); bkhan66688@gmail.com (A. 4); luca.marconi@socialthingum.com (A. 5); ricardo.matamoros@socialthingum.com (A. 6)
ORCID: 0000-0002-4828-4239 (A. 1); 0000-0003-1839-9366 (A. 2); 0000-0002-1850-5846 (A. 3); [0000-0001-8567-4298 (A. 4); 0000−0002−0236−6159 (A. 5); 0000−0002−1957−2530 (A. 6)

prediction [7, 22]. Different techniques and algorithms are used throughout the process, with deep learning seeing the highest usage rates [8, 23]. In this study, propose a deep learning framework to predict the gender and age group of face images with a high accuracy rate. This framework is built on the ensemble of attention and residual Convolutional networks.

The aims of the proposed model are to use the approach of deep learning technique in the prediction process, the effective accuracy and predictive mining and evaluate it reaching to get best results of prediction and overcome the problems of time, accuracy and processing load. Section 2 presents the related work while Section 3 discusses the proposed methodology. The Section 4 elaborate the results of proposed model while Section 5 conclude the paper.

2. Related Work

Prediction of age and gender from the face photos, as a special problem of face analysis has been attracting attention in recent years [9]. There are many works done so far in the prediction of age and gender from facial images. Here are review of the most promising research work.

Nada et al., [10] conducted research on validation and prediction of gender and age using CNN for one single image. The UP-student's dataset was used in the experiment to evaluate the suggested method. Sadly, age estimation deteriorated due to the pro-posed solution's poor gender prediction performance [24]. Overall, both genders had a gender prediction accuracy of roughly 82%. Additionally, the algorithm performs better when guessing images of male faces (89%, compared to 74% for females). After examining the photographs where the model failed to correctly estimate the gender, there were a number of causes. The primary cause at some ages, the distinction between the facial characteristics of men and women is not always as obvious as it ought to be. Hijab also conceals various facial characteristics in photographs of women. Finally, there is a flaw in the model that was utilized; it did not assign the moustache enough importance in predicting gender. Considering the age prediction findings, it was not so good that the total forecast accuracy for both genders was just 57%.

Al-Azzawi, [11] used Adience Benchmark dataset of face images; it consists of 17603 images of human faces for variance of ages and genders. The ages of the persons in the dataset are classified into 10 groups and the gender binary is classified into two types. The images of the datasets are divided into two sets equally, one for the training phase and another for the testing phase. They used Mean Absolute Error (MAE) function to evaluate the implementation of age prediction, and the accuracy of the gender prediction assessed by the hit ratio to compare the current proposed Deep Multi-tasking CNN [25]. The accuracy for gender detection was (91%) while accuracy in mean absolute error (MAE) for age prediction was (4.00) in CNN and DMTL model.

3. Proposed Method

This section will discuss the design and components used in the proposed system. The dataset and its features are also discussed in this section.

3.1. Proposed Model

The proposed system is being created and developed while keeping in mind all of the shortcomings and restrictions of the current system, thus anticipate that it will be an acceptable system that successfully satisfies all of the goals of the current system.

The reasons behind choosing of this model are to improve accuracy, for both of the classifiers i.e. age and gender, to improve training time for both classifier, and to minimize the resource utilization of computer. A new CNN layer architecture is de-signed which works much better than other Built-in models like VGG, resnet50 and Mobile net etc. Especially for that particular problem. It contains the Convolution layer, Batch Normalization layer, Max and average pooling (at the last) and also dropout layers along with fully connected layers. Here it used ReLU, Elu and Soft-max activation functions. The model takes input layer of image size= (128, 128, 3) and used 3x3 kernel/filter. Also used bias constraints, filter of 64 to 512 and stride of (1, 1) with same padding. Figure 1 illustrates the proposed model architecture.

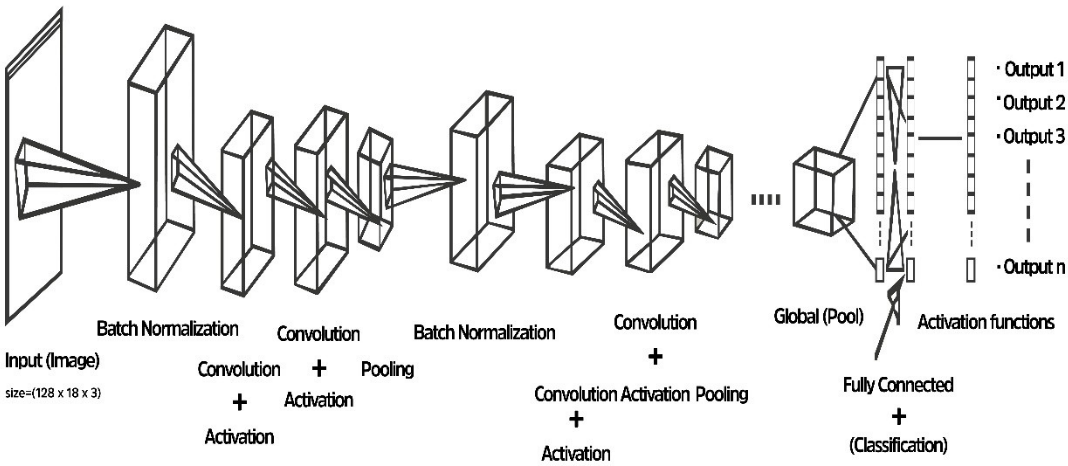

Figure 1: Proposed model architecture

3.2. Data Set

The Adience dataset served as the basis for this study. The fundamental tenet of the data collection is to record the photographs as accurately as possible, taking into account any variances in look, posture, lighting, and image quality, to mention a few. Almost 26000 pictures were used along with labels of age and gender group. The Adience benchmark dataset, which consists of face photographs that are automatically posted from smart phones to Flickr, is made for age and gender categorization. These photos show unfiltered, unedited photographs from the real world and social media. So, need to adjust them according to suitable way by performing preprocessing.

3.3.Structure of Proposed Model

A single Convolutional neural network with two heads (output branches) is utilized in this multi-task learning issue to simultaneously predict age and gender with the following properties.

- Number of Epoch for both Age and Gender module is 12
- Size of image is 128 x 128.
- Total number of parameters 6059022.
- Batch size used is 64.
- Time required for both classifiers is 45 min Approx.
- “Sparse categorical Cross entropy” as loss functions.
- 20% data for testing and 80% for training.
- 25 deep layers.

3.4.Training and Testing

In the training and testing phases of the current network, Adience Benchmark dataset of face images was used; it consists of Approx. 26000 images of human faces for variance of ages and genders. The ages of the persons in the dataset are classified into 8 groups and the gender binary is classified into two types. The images of the datasets are divided into two sets equally, one for the training phase and another for the testing phase. The image size 128x128 was used in the training and testing phase. The Epoch for gender is 15 and Epoch for Age is 12, and the time taken for both classifier to train is 90 min (45 min each).

4. Results and Analysis

The gender classification model is tested and evaluated using machine learning evaluation metrics. The below section will discuss the results of gender classification model.

4.1. Gender Classification model

Gender prediction is viewed as a classification issue, and this network's output layer is a Softmax with two nodes that represent the classifications of male and female. The model is implemented as a network with three layers, two of which are output layers and one of which is a fully linked layer. The anticipated values for each class may be obtained from the gender prediction network by loading this model into memory and sending the output of the face detection process (detected face) through the network. Now that the output has reached its maximum value, that may utilize that number to determine a person's gender. Table 1 presents the classification report of gender model while Figure 2 illustrates the accuracy comparison of gender model with other models.

Table 1
Classification Report for the Gender Model

(Classes)	(Precision)	(Recall)	(F1-Score)	(Support)
Female (0)	0.92	0.92	0.92	2744
Male (1)	0.92	0.91	0.91	2492
Accuracy			0.95	5236

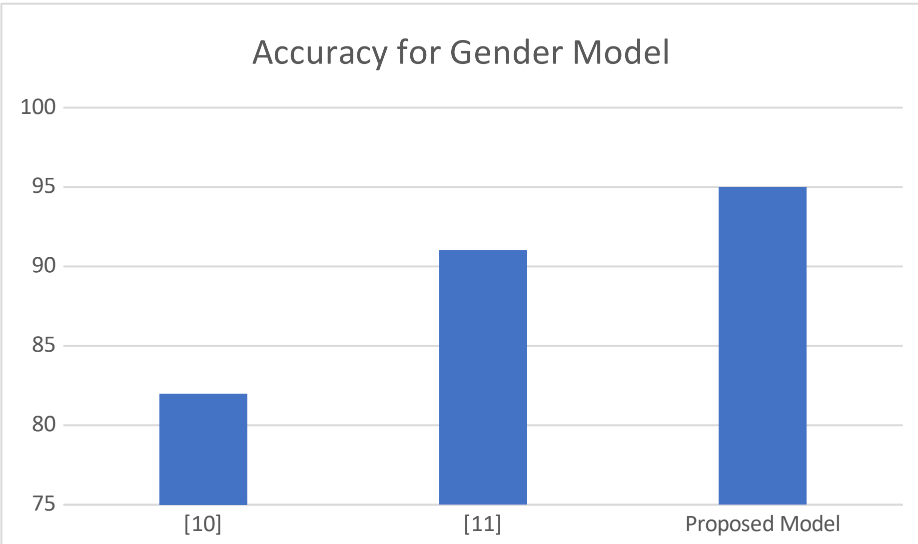

Figure 2: Accuracy comparison between gender models

4.2. Age Classification model

People find it exceedingly difficult to make accurate age predictions by simply looking at a person, but it is possible to do so when estimating a range of ages. Consequently, regarded it as a classification issue using the Adience Benchmark dataset. The predicted values for all training may be obtained from the network by loading this model into memory and running the output of the face detection process (detected face) through the age prediction network. As, utilized the output's maximum value as a forecast age group by taking that value. Table 2 presents the classification report of age model while Figure 3 illustrates the accuracy comparison of age model with other models

Table 2
Classification Report for Age Model

(Classes)	(Precision)	(Recall)	(F1-Score)	(Support)
(0-2)	0.83	0.83	0.83	422
(4-6)	0.69	0.85	0.76	625
(8-13)	0.86	0.63	0.73	662
(15-20)	0.61	0.56	0.59	558
(25-32)	0.71	0.79	0.75	1558
(38-43)	0.59	0.65	0.62	879
(48-53)	0.53	0.38	0.44	267
(60+)	0.91	0.55	0.69	265
Accuracy		0.92	5236	

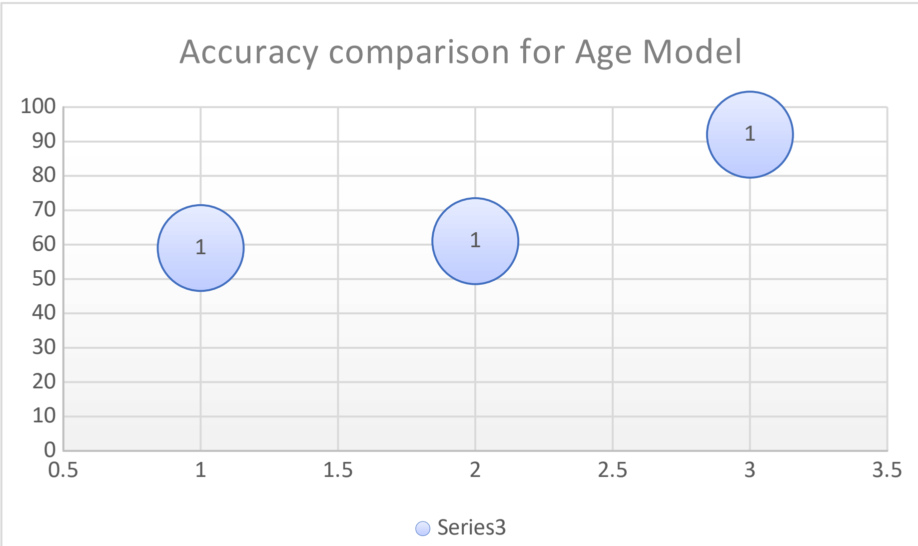

Figure 3: Accuracy comparison of Age Model

The proposed model gives 95% Accuracy for Gender Classifier and 92% for Age Classifier which is highest among all the previous/related work discussed. The model takes much short time to train up for both of the classifier i.e. Age and Gender Classifier. model utilizes the computing resources (RAM, CPU, and GPU) in much optimized manner and the computing cost is also lower for the model as compared to the previous model discussed in existing model session.

5. Conclusion

Age, gender, and the age range of a person's personal photo have recently become crucial pieces of information for many businesses and governments to use for commercial, identity, security, and other purposes [12]. Additionally, because this information was gathered from people using an enterprise system, form validation was suggested as a way to lower user data entry mistakes. The module is mainly designed for biometrics research in social applications for the future where the content is to be shown for some specific gender and age group and it must be possible to predict and disclose information about each person [13, 26]. The experimental investigation found that the suggested CNNs had a fair classification accuracy after being trained quickly with a large number of photos. The proposed CNNs will be utilized in next work for social media statistics and gender categorization in mobile applications. Secondly, the model is just trained on Adience benchmark dataset that means more sophisticated systems can use more training data. It is possible that the results can be significantly improved in future beyond the results reported here.

6. References

[1] Hizam, S. M., Ahmed, W., Fahad, M., Akter, H., Sentosa, I., & Ali, J. User behavior assessment towards biometric facial recognition system: A SEM-neural network approach. In Future of Information and Communication Conference (pp. 1037-1050). Springer, Cham (2021).

[2] Zhu, J., Luo, B., Zhao, S., Ying, S., Zhao, X., & Gao, Y. IExpressNet: Facial Expression Recognition with Incremental Classes. In Proceedings of the 28th ACM International Conference on Multimedia (pp. 2899-2908) (2020).

[3] Shinde, G. R., Kalamkar, A. B., Mahalle, P. N., Dey, N., Chaki, J., & Hassanien, A. E. Forecasting models for coronavirus disease (COVID-19): a survey of the state-of-the-art. SN Computer Science, 1(4), 1-15 (2020).

[4] Chekroud, A. M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., ... & Choi, K. The promise of machine learning in predicting treatment outcomes in psychiatry. World Psychiatry, 20(2), 154-170 (2021).

[5] Maior, C. B. S., das Chagas Moura, M. J., Santana, J. M. M., & Lins, I. D. Real-time classification for autonomous drowsiness detection using eye aspect ratio. Expert Systems with Applications, 158, 113505 (2020).

[6] Ma, C., Wei, J. L., Liu, B., Ding, M., Yuan, L., Han, Z., & Poor, H. V. Trusted AI in Multi-agent Systems: An Overview of Privacy and Security for Distributed Learning. arXiv preprint arXiv:2202.09027 (2022).

[7] Alonso-Fernandez, F., Hernandez-Diaz, K., Ramis, S., Perales, F. J., & Bigun, J. Facial masks and soft-biometrics: Leveraging face recognition CNNs for age and gender prediction on mobile ocular images. IET Biometrics, 10(5), 562-580 (2021).

[8] Tufek, N., Yalcin, M., Altintas, M., Kalaoglu, F., Li, Y., & Bahadir, S. K. Human action recognition using deep learning methods on limited sensory data. IEEE Sensors Journal, 20(6), 3101-3112 (2019).

[9] Carletti, V., Greco, A., Percannella, G., & Vento, M. Age from faces in the deep learning revolution. IEEE transactions on pattern analysis and machine intelligence, 42(9), 2113-2132 (2019).

[10] Abu Nada, A.M., Alajrami, E., Al-Saqqa, A.A. and Abu-Naser, S.S., Age and Gender Prediction and Validation Through Single User Images Using CNN (2020).

[11] Al-Azzawi, D.S,. Human Age and Gender Prediction Using Deep Multi-Task Convolutional Neural Network. Journal of Southwest Jiaotong University, 54(4) (2019).

[12] Zeebaree, S., Ameen, S., & Sadeeq, M. Social media networks security threats, risks and recommendation: A case study in the kurdistan region. International Journal of Innovation, Creativity and Change, 13, 349-365 (2020).

[13] Faundez-Zanuy, M., Fierrez, J., Ferrer, M. A., Diaz, M., Tolosana, R., & Plamondon, R. Handwriting biometrics: Applications and future trends in e-security and e-health. Cognitive Computation, 12(5), 940-953 (2020).

[14] Bosman, L., Fernhaber, S., & SpringerLink (Online service). Teaching the entrepreneurial mindset to engineers. Switzerland: Springer International Publishing (2018).

[15] Othmani, A., Taleb, A. R., Abdelkawy, H., & Hadid, A. Age estimation from faces using deep learning: A comparative analysis. Computer Vision and Image Understanding, 196, 102961 (2020).

[16] Korban, M., Youngs, P., & Acton, S. T. TAA-GCN: A Temporally Aware Adaptive Graph Convolutional Network for Age Estimation. Pattern Recognition, 109066 (2022).

[17] Sun, N., Zhang, J., Rimba, P., Gao, S., Zhang, L. Y., & Xiang, Y. Data-driven cybersecurity incident prediction: A survey. IEEE communications surveys & tutorials, 21(2), 1744-1772 (2018).

[18] Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N. Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195-204 (2019).

[19] Hosny, A., Parmar, C., Coroller, T. P., Grossmann, P., Zeleznik, R., Kumar, A., ... & Aerts, H. J. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. PLoS medicine, 15(11), e1002711 (2018).

[20] Tahir, R., Batool, B., Jamshed, H., Jameel, M., Anwar, M., Ahmed, F., ... & Zaffar, M. F. Seeing is believing: Exploring perceptual differences in deepfake videos. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1-16) (2021, May).

[21] Khanagar, S. B., Vishwanathaiah, S., Naik, S., Al-Kheraif, A. A., Divakar, D. D., Sarode, S. C., ... & Patil, S. Application and performance of artificial intelligence technology in forensic odontology–A systematic review. Legal Medicine, 48, 101826 (2021).

[22] Benkaddour, M. K., Lahlali, S., & Trabelsi, M. Human Age and Gender Classification using Convolutional Neural Network. In 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH) (pp. 215-220). IEEE (2021, February).

[23] Ghosh, P., Azam, S., Karim, A., Jonkman, M., & Hasan, M. Z. Use of efficient machine learning techniques in the identification of patients with heart diseases. In 2021 the 5th International Conference on Information System and Data Mining (pp. 14-20) (2021, May).

[24] Rai Jain, P., MK Quadri, S., & Lalit, M. Recent Trends in Artificial Intelligence for Emotion Detection using Facial Image Analysis. In 2021 Thirteenth International Conference on Contemporary Computing (IC3-2021) (pp. 18-36) (2021, August).

[25] Ghosh, P. Deep Learning to Diagnose Diseases and Security in 5G Healthcare Informatics. In Machine Learning and Deep Learning Techniques for Medical Science (pp. 279-331). CRC Press (2022).

[26] Shopon, M., Hossain Bari, A. S. M., Bhatia, Y., Narayanaswamy, P. K., Tumpa, S. N., Sieu, B., & Gavrilova, M. Biometric System De-identification: Concepts, Applications, and Open Problems. In Handbook of Artificial Intelligence in Healthcare (pp. 393-422). Springer, Cham (2022).