2019 International Conference on Innovative Computing (ICIC)

Age and Gender Prediction using Deep Convolutional Neural Networks

Insha Rafique
Department of Software Engineering
University of Management and Technology
Lahore, Pakistan
insharafique05@gmail.com

Awais Hamid
Department of Software Engineering
University of Management and Technology
Lahore, Pakistan
awaishamid555@gmail.com

Sheraz Naseer
Department of Computer Science & Engineering
University of Engineering and Technology, Lahore, Pakistan
sheraz.naseer@gmail.com

Muhammad Asad
Department of Software Engineering
University of Management and Technology
Lahore, Pakistan
asaddhollaumt@gmail.com

Muhammad Awais
Department of Software Engineering
University of Management and Technology
Lahore, Pakistan
awaismalik3577@gmail.com

Talha Yasir
Department of Computer Science
University of Management and Technology
Lahore, Pakistan
talha.yasirusmani@gmail.com

***Abstract*—Age and gender identification have become a major part of the network, security and care. It has a common use in age specific content access for children. Social media uses it in delivering layered ads and marketing to extend it's a reach. Face recognition has developed to a great extent that we have to map it further in getting more useful results having different approaches. In this paper, we propose deep CNN to improve age and gender predication from significant results can be obtained and a significant improvement can be seen in various tasks such as face recognition. A simple convolutional network architecture is proposed to make a noticeable improvement in this field using existing methods. Using deep CNN, model is trained to an extent that accuracy of Age and Gender become 79% using HAAR Feature-based Cascade Classifiers is an effective method proposed by Paul Viola and Michael Jones. It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images.**

***Keywords—** Age and Gender prediction, Deep Convolutional Neural Networks, Deep learning, CNN*

I. INTRODUCTION

Age and gender, two of the key facemask attributes, play a very initial role in social communications, making age and gender approximation from a single image an important task in intelligent applications, such as access control, human-computer interaction, law application, marketing intelligence and visual observation, etc. It can be used to suppose the age and gender of the user and use this information to make modified product and understanding for each user. It plays the vital role in marketing for the marketer by addressing the target audience on the basis of age and gender.

Age recognition plays a major role in Police investigation and Intelligence department as it is helpful in finding the actual suspect on the basis of his age. They could get a filtered out result of that person who has performed criminal act or any other activity. If a person gives a biased opinion about his age after getting result from a age recognition software then the actual age and predicted age would be approximately same Which tells its reliability and this reliability make a trust factor for many other useful operations in daily life.

Speech and text vary from gender to gender and also with age. Prediction of Age and Gender still lacks its accuracy those move efficient way of predication is on the basis of face so that it can needs of commercial applications. Most of techniques presented in literature uses face recognition for prediction of age and gender but they use feature dimensions of face which is not as much efficient. Classification schemes are used for this purpose. Moreover machine learning also didn't exploit a large number of images and data from various sources through internet for improvement of classification activities [1].

In this paper we purpose a scheme to fill the gap between automatic face recognition and age and gender prediction. In the past when there is an improvement done on face recognition on large scale, at that point a link between face recognition and Convolution Neural Network (CNN) is proposed and by studying it further we created a system in which a limited amount of face data sets are used to accurately predict age and gender [2]. A data of unfilled images is taken, despite of complexity we have in network design it performs well and gives a cutting edge result. It provides a base to deep learning and suggests us that there is still a lot of room for improvements and the mysteries still remains unsolved. A data of pertained images is taken that are used to train model using HAAR Feature-based Cascade Classifiers [3].

Our objective is to make a system that would be efficient enough to predict age and gender of a person without breaching his security and other bypassing any other security. Such efficient systems are helpful in variety of ways in performing different activities. Our main objective is to train a model which can predict age and gender in most efficient way. It is also aimed to use in age specific content access limitation by which system can detect age and gender and allows/deny user to access that content [4].

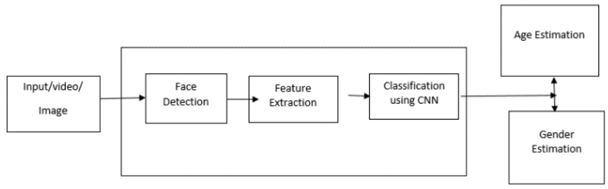

Fig 1: Age and Gender Model

The relevant applications of age and gender predictions systems are growing rapidly in recent days due its important modules and beneficial uses for computer vision application. There are some use cases that demonstrate the problem and we are going to address these problems [5].

Moreover the best use of age and gender recognition is in social media, marketing and advertisements [6]. Social media platforms can predict your age and genders and shows such contents which is of interest level of people of that age limit. It also helps growing e-commerce business and internet marketing business as it would help them advertise products according to interest level of that age group which in turn gives benefit to users too [7].

II. RELATED WORK

The problem of automated age extracting attributes have been in attention for a long period of time. Early classification of age was done by calculation ration between different features of face like nose, eyes, mouth, chin etc. After localizing calculating their sizes and distances, ratio between them are calculated in order to predict age by using conventional methods [8]. Recently a model [9] is proposed to show age progression of people under 18 years old but that doesn't works and pictures are different on social networks.

There are some other methods that predicts age as manifold [10] But it requires well aligned and front facing images exactly, thus that proposition also don't meets the needs. So these methods gives only experimental result on limited data sets. Thus such methods are inappropriate to use in independent datasets [11].

There is another method in which distribution of facial patches is discussed which is known is Gaussian Mixture Model (GMM) [12]. It is used for the representation of distribution of local facial measurements but is this instead of pixel matches, robust image matching algorithm is used. So we cannot suggest it good model as we need to find out face patch distribution and for this Hidden-Markov-Model can be used.

Robust image descriptor technique is an alternate way for local image intensity patches [13]. Gabor image descriptor method is used in used along with Fuzzy-LDA classifier to detect image of face belonging to more than one age group [14]. As a whole Biological Inspired features (BIF) [15] and other manifold-learning features are used for age prediction. Gabor image descriptor and local binary pattern (LBP) were used with hierarchical age classifier method consists of Support Vector Machines (SVM) [16]are used to classify input image to a specific age class by using support vector regression so that the result obtained would be precise [17].

All of the above proposed models are used to have a precise age result by using different techniques and methods that are described in such models and performs a benchmark in age estimation. The best proposed model was implemented on a group photo having multiple faces in it. On this photo we implemented local binary pattern (LSB) descriptor and Support Vector Machine (SVM) classifier.

There are various method to classify gender. One of the method of gender classification is by using neural networks trained on the small set of frontal faces image [18]. It uses 3D structure of head and image intensities for classifying age.

Support vector machine (SVM) classifier were used and applied to image intensities. In spite of using SVM, we can use Ada boost which is a replacement of SVM having same functionalities.

In the recent, Weber's Local texture descriptor is used which gives a perfect result on Face Recognition Technology (FERET). In this concentration, form and quality structures used to predict the most perfect and accurate result on FERET benchmark. Most of the methods described above uses FERET benchmark to develop a most precise system giving accurate results [19]. FERET images were taken to extremely measured complaint and the result obtained from them are highly saturated. It is actually difficult to find out actual advantages of these techniques. Face recognition is also used on Labelled faces in the wild (LFW) and this method has a combination of Local binary pattern (LBP) and an Ada boost classifier [20].

In the past when computers were not efficient then it was a great job to make computer perform tasks as the interaction of people vary to each computer. Every person cannot respond to computer in similar way [21]. There was need of making Human Computer Interaction (HCI) easy so that it would be accessible for everybody. But now due to age predictor computer can understand how to respond to a person of that age.

Thus the methods presented above, we have noticed that we used dataset which is more difficult and challenging them images used in Labelled faces in the wild (LFW) and gives more accurate results in robust system and gets more information form massive data set [7].

III. METHODOLOGY

The first application of Convolutional Neural Network (CNN) is LeNet-5 network [22] by using optical character recognition. If we compare this activity with modern deep convolutional network technique it is considered to be very simple and humble as that time there were limited computational resources and there are challenges to train algorithms [23].

Now the time has come when neural networks become so deep that they became prevalent due to increase in computational resources and the training data is easily available on internet. Moreover now such methods are available that can train data easily and readily. Now there are various application of Convolutional Neural Network (CNN) are present like human pose estimation, face parsing, facial key point detection, and speech recognition and action

Classification. On unconstraint photo this is their first application according to our knowledge.

We have noticed that if we want to gather large datasets of images from social platform them it may require their privacy permission or may become a security hazard and its very time taking to label is manually [24]. Dataset from real world social images we have noticed that they are limited in size and they have no match in size with larges database image sets.

Over fitting is a common problem while using machine learning based methods on small image collections [25]. This problem is intensified when considering deep convolutional neural networks due to large parameters. So we have to be very careful while using such methods.

The system we proposed works perfectly fine with experiments in classification for age and gender. Our networks consist of three convolutional layers in which two of them are fully connected with small number of neurons. We use small network design for taking less risk for over fitting and also for the nature of problem we are going to solve. Classification of age on dataset requires to differentiate between eight classes and for two genders. Thus we can say that ten thousand classes are used to train the datasets used for face recognition.

The retail store owner need the analytics of his store. He wants to find the number of visitors by age group and genders. He also wants to find the most popular aisles and the articles viewed by age group [26].
The system needs in smoking areas for age detection to stop the growth of teenage smokers. The smoking areas should ahead the cameras to identify the age group of the smokers.

TABLE 1. AGE AND GENDER CLASSES

Age and Gender Group	15-20	25-32	38-43	48-53	60+	Total
Male	734	2308	1294	392	442	8192
Female	919	2589	1056	433	427	9411

IV. EXPERIMENTAL WORK

We trained our technique on the as of late proposed informational index of countenances for age and sex order. The data-set of countenances contains naturally transferred Flickr pictures. As the pictures were naturally transferred without earlier sifting, they portray Testing in-the-wild settings and differ in outward appearance, head present, impediments, lighting conditions, picture quality and so on. Additionally, a portion of the pictures are of exceptionally low quality or contain extraordinary movement obscure.

The figure above (first figure in the post) illustrates example images from the dataset of Faces. Below is a breakdown of the dataset into the different age and gender classes [27]. A number of other prediction models are proposed in literature and can be found in [28-41].

A. *Initialization*

The weight in all the layers are initialized with random values to zero which means Gaussian with standard deviation of 0.01. We do not use models that are already trained but we trained them from scratch without using any outside labelled datasets.

V. ARCHITECTURE OF CNN

Five network architecture layers are used in this model. Two of them are fully connected layers and three convolutional layers are used.

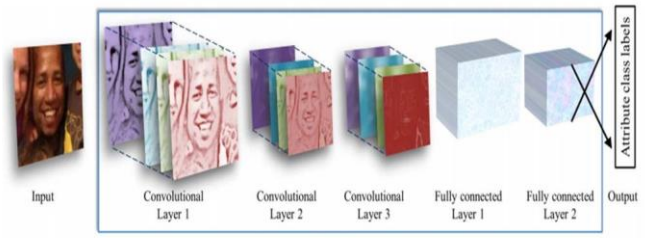

Fig 2: Architecture of used CNN

VI. RESULTS AND EVALUATIONS

We tried different things with two techniques for characterization:

- Centre Crop: Feeding the system with the face picture trimmed to 227 × 227 around the face focus [22].

- Over-testing: We separate five 227 × 227 pixel crop locales, four from the edges of the 256 × 256 face picture and one from the focal point of the face alongside their flat flips [23]. Each of the 10 yields are encouraged to the system and the last order is the normal of the forecasts of the 10 crops [24].

The tables below summarize our results compared to previously proposed methods.

TABLE 2. Gender Predication

Method	**Exact**	**1-off**
Best [28]	45.1 ± 2.6	79.5 ±1.4
Proposed by single produce	49.5 ± 4.4	84.6 ± 1.7
Proposed by over-sampling	**50.7 ± 5.1**	**84.7 ± 2.2**

We measure mean exactness + standard variety, 1-off in age arrangement implies the age expectation was either right or 1-off from the right age class.

TABLE 3. AGE PREDICATION

Method	Accuracy
Best [2]	77.8 ± 1.3
Best [29]	79.3 ± 0.0
Proposed by single produce	85.9 ± 1.4
Proposed by over-sampling	86.8 ± 1.4

VII. CONCLUSION

In this article age and gender detection system with artificial neural network is proposed and implemented. The age and gender estimation system consist of face detection and model training for the classification. Using deep CNN, model is trained to an extent that accuracy of Age and Gender become 79% using HAAR cascading. Its accuracy could be increased more using more efficient algorithms and more precise architecture of CNN so that it could have been used more in different platforms.

REFERENCES

[1] A. A. Zaidan, B. B. Zaidan, A. Al-Haiqi, M. L. M. Kiah, M. Hussain, and M. Abdulnabi, "Evaluation and selection of open-source EMR software packages based on integrated AHP and TOPSIS," *J. Biomed. Inform.*, vol. 53, pp. 390–404, 2015.

[2] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 12, pp. 2037–2041, 2006.

[3] R. Sharma, T. S. Ashwin, and R. M. R. Guddeti, "A Novel Real-Time Face Detection System Using Modified Affine Transformation and Haar Cascades," in *Recent Findings in Intelligent Computing Techniques*, Springer, 2019, pp. 193–204.

[4] M. Hussain, A. Al-Haiqi, A. A. Zaidan, B. B. Zaidan, M. L. M. Kiah, N. B. Anuar, and M. Abdulnabi, "The landscape of research on smartphone medical apps: Coherent taxonomy, motivations, open challenges and recommendations," *Comput. Methods Programs Biomed.*, vol. 122, no. 3, pp. 393–408, 2015.

[5] N. Kalid, A. A. Zaidan, B. B. Zaidan, O. H. Salman, M. Hashim, and H. Muzammil, "Based real time remote health monitoring systems: A review on patients prioritization and related" big data" using body sensors information and communication technology," *J. Med. Syst.*, vol. 42, no. 2, p. 30, 2018.

[6] M. Uricár, R. Timofte, R. Rothe, J. Matas, and L. Van Gool, "Structured output svm prediction of apparent age, gender and smile from deep features," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 25–33.

[7] B. B. Zaidan, A. Haiqi, A. A. Zaidan, M. Abdulnabi, M. L. M. Kiah, and H. Muzamel, "A security framework for nationwide health information exchange based on telehealth strategy," *J. Med. Syst.*, vol. 39, no. 5, p. 51, 2015.

[8] G. Ozbulak, Y. Aytar, and H. K. Ekenel, "How transferable are CNN-based features for age and gender classification?," in *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2016, pp. 1–6.

[9] H. Yang, D. Huang, Y. Wang, and A. K. Jain, "Learning Continuous Face Age Progression: A Pyramid of GANs," *arXiv Prepr. arXiv1901.07528*, 2019.

[10] D. Sabbagh, P. Ablin, G. Varoquaux, A. Gramfort, and D. A. Engeman, "Manifold-regression to predict from MEG/EEG brain signals without source modeling," *arXiv Prepr. arXiv1906.02687*, 2019.

[11] M. Abdulnabi, A. Al-Haiqi, M. L. M. Kiah, A. A. Zaidan, B. B. Zaidan, and M. Hussain, "A distributed framework for health information exchange using smartphone technologies," *J. Biomed. Inform.*, vol. 69, pp. 230–250, 2017.

[12] K. Vaibhav, J. Prasad, and B. Singh, "Convolutional Neural Network for Classification for Indian Jewellery," *Available SSRN 3351805*, 2019.

[13] Q. Shen, G. Xiao, Y. Zheng, J. Wang, Y. Liu, X. Zhu, F. Jia, P. Su, B. Nie, and F. Xu, "ARMBIS: Accurate and Robust Matching of Brain Image Sequences from Multiple Modal Imaging Techniques," *Bioinformatics*, 2019.

[14] M. Vasileiadis, G. Stavropoulos, and D. Tzovaras, "Facial Soft Biometrics Detection on Low Power Devices," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, p. 0.

[15] F. Dornaika, I. Arganda-Carreras, and C. Belver, "Age estimation in facial images through transfer learning," *Mach. Vis. Appl.*, vol. 30, no. 1, pp. 177–187, 2019.

[16] *University of California Regents v. Bakke BT - US*, vol. 438, no. No. 76-811. Supreme Court, 1978, p. 265.

[17] M. Hussain, A. Al-Haiqi, A. A. Zaidan, B. B. Zaidan, M. Kiah, S. Iqbal, S. Iqbal, and M. Abdulnabi, "A security framework for mHealth apps on Android platform," *Comput. Secur.*, vol. 75, pp. 191–217, 2018.

[18] S. Hickson, N. Dufour, A. Sud, V. Kwatra, and I. Essa,

"Eyemotion: Classifying facial expressions in VR using eye-tracking cameras," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2019, pp. 1626–1635.

[19] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2015, pp. 34–42.

[20] M. Hussain, A. A. Zaidan, B. B. Zidan, S. Iqbal, M. M. Ahmed, O. S. Albahri, and A. S. Albahri, "Conceptual framework for the security of mobile health applications on android platform," *Telemat. Informatics*, vol. 35, no. 5, pp. 1335–1354, 2018.

[21] M. Afifi and A. Abdelhamed, "AFIF4: deep gender classification based on adaboost-based fusion of isolated facial features and foggy faces," *J. Vis. Commun. Image Represent.*, vol. 62, pp. 77–86, 2019.

[22] T. Li, D. Jin, C. Du, X. Cao, H. Chen, J. Yan, N. Chen, Z. Chen, Z. Feng, and S. Liu, "The image-based analysis and classification of urine sediments using a LeNet-5 neural network," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, pp. 1–6, 2019.

[23] G. Antipov, M. Baccouche, S.-A. Berrani, and J.-L. Dugelay, "Effective training of convolutional neural networks for face-based gender and age prediction," *Pattern Recognit.*, vol. 72, pp. 15–26, 2017.

[24] B. A. Jonsson, G. Bjornsdottir, T. E. Thorgeirsson, L. M. Ellingsen, G. B. Walters, D. F. Gudbjartsson, H. Stefansson, K. Stefansson, and M. O. Ulfarsson, "Deep learning based brain age prediction uncovers associated sequence variants," *bioRxiv*, p. 595801, 2019.

[25] S. Chen, C. Zhang, M. Dong, J. Le, and M. Rao, "Using ranking-cnn for age estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5183–5192.

[26] V. I. Iglovikov, A. Rakhlin, A. A. Kalinin, and A. A. Shvets, "Paediatric Bone age assessment using deep convolutional neural networks," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 2018, pp. 300–308.

[27] R. Rothe, R. Timofte, and L. Van Gool, "Deep expectation of real and apparent age from a single image without facial landmarks," *Int. J. Comput. Vis.*, vol. 126, no. 2–4, pp. 144–157, 2018.

[28] E. Eidinger, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces," *IEEE Trans. Inf. Forensics Secur.*, vol. 9, no. 12, pp. 2170–2179, 2014.

[29] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2234–2240, 2007.

[30] A. A. Zaidan, B. B. Zaidan, A. Al-Haiqi, M. L. M. Kiah, M. Hussain, and M. Abdulnabi, "Evaluation and selection of open-source EMR software packages based on integrated AHP and TOPSIS," *J. Biomed. Inform.*, vol. 53, pp. 390–404, 2015.

[31] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 12, pp. 2037–2041, 2006.

[32] R. Sharma, T. S. Ashwin, and R. M. R. Guddeti, "A Novel Real-Time Face Detection System Using Modified Affine Transformation and Haar Cascades," in *Recent Findings in Intelligent Computing Techniques*, Springer, 2019, pp. 193–204.

[33] M. Hussain, A. Al-Haiqi, A. A. Zaidan, B. B. Zaidan, M. L. M. Kiah, N. B. Anuar, and M. Abdulnabi, "The landscape of research on smartphone medical apps: Coherent taxonomy, motivations, open challenges and recommendations," *Comput. Methods Programs Biomed.*, vol. 122, no. 3, pp. 393–408, 2015.

[34] N. Kalid, A. A. Zaidan, B. B. Zaidan, O. H. Salman, M. Hashim, and H. Muzammil, "Based real time remote health monitoring systems: A review on patients prioritization and related" big data" using body sensors information and communication technology," *J. Med. Syst.*, vol. 42, no. 2, p. 30, 2018.

[35] M. Uricár, R. Timofte, R. Rothe, J. Matas, and L. Van Gool, "Structured output svm prediction of apparent age, gender and smile from deep features," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 25–33.

[36] B. B. Zaidan, A. Haiqi, A. A. Zaidan, M. Abdulnabi, M. L. M. Kiah, and H. Muzamel, "A security framework for nationwide health information exchange based on telehealth strategy," *J. Med. Syst.*, vol. 39, no. 5, p. 51, 2015.

[37] G. Ozbulak, Y. Aytar, and H. K. Ekenel, "How transferable are CNN-based features for age and gender classification?," in *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2016, pp. 1–6.

[38] H. Yang, D. Huang, Y. Wang, and A. K. Jain, "Learning Continuous Face Age Progression: A Pyramid of GANs," *arXiv Prepr. arXiv1901.07528*,

[39] N. Kalid, A. A. Zaidan, B. B. Zaidan, O. H. Salman, M. Hashim, and H. Muzammil, "Based real time remote health monitoring systems: A review on patients prioritization and related" big data" using body sensors information and communication technology," *J. Med. Syst.*, vol. 42, no. 2, p. 30, 2018.

[40] Y. Zhang, “Age, gender, and Internet attitudes among employees in the business world,” *Comput. Human Behav.*, vol. 21, no. 1, pp. 1–10, 2005.

[41] D. Sabbagh, P. Ablin, G. Varoquaux, A. Gramfort,and D. A. Engeman, “Manifold-regression to predict from MEG/EEG brain signals without source modeling,” *arXiv Prepr. arXiv1906.02687*, 2019.