



International Journal of Artificial Intelligence and Soft Computing

ISSN online: 1755-4969 - ISSN print: 1755-4950

<https://www.inderscience.com/ijaisc>

Age and gender prediction using Haar cascade algorithm and fine-tuned CNN framework

Danish Ali, Atta Ur Rahman, Bibi Saqia, Maqbool Khan

DOI: [10.1504/IJAISC.2025.10072771](https://doi.org/10.1504/IJAISC.2025.10072771)

Article History:

Received:	13 April 2024
Last revised:	22 September 2024
Accepted:	29 October 2024
Published online:	25 August 2025

Age and gender prediction using Haar cascade algorithm and fine-tuned CNN framework

Danish Ali* and Atta Ur Rahman

Riphah Institute of System Engineering (RISE),
Riphah International University Islamabad,
46000, Pakistan

Email: danish.ali@riphah.edu.pk

Email: atta.rahman@riphah.edu.pk

*Corresponding author

Bibi Saqia

Department of Computer Science,
University of Science and Technology,
Bannu, 28100, Pakistan
Email: b3lt.k.060@gmail.com

Maqbool Khan

Pak-Austria Fachhochschule: Institute of Applied
Sciences and Technology,
Haripur, Pakistan
Email: maqbool.khan@fecid.pafias.edu.pk

Abstract: Age and gender prediction using unfiltered facial images is an essential task these days. We apply deep learning algorithm convolutional neural network (CNN) with multiple datasets such as UTKFace, Adience, and two custom datasets all combined for best performance. When we glanced over the existing works, we found that the existing models are efficient regarding age prediction, another problem was the computational power requirement, which is less useful in real-time and without a GPU. In this research, we used fewer layers to make this model usable in real-time and eliminate the requirement of GPUs for computational power. We employed the Haar cascade algorithm for face detection, and CNN with 12 layers for the other two modules, age prediction, and gender prediction. Combining different datasets and training each module on different balanced and unbiased datasets, we have achieved a higher accuracy regarding age prediction.

Keywords: age; gender; convolutional neural network; CNN; Haar cascade algorithm; UTKFace.

Reference to this paper should be made as follows: Ali, D., Rahman, A.U., Saqia, B. and Khan, M. (2025) 'Age and gender prediction using Haar cascade algorithm and fine-tuned CNN framework', *Int. J. Artificial Intelligence and Soft Computing*, Vol. 9, No. 1, pp.22–34.

Biographical notes: Danish Ali received his BS (Hons) degree from Riphah International University Islamabad Pakistan. Currently, his MS is in progress at Riphah International University Islamabad Pakistan.

Atta Ur Rahman received his BS (Hons) in Telecommunication, and MS in Computer Science from the University of Science and Technology Bannu in 2018, and PhD in Computer Science from Ghulam Ishaq Khan Institute of Engineering Sciences and Technology in 2022. Currently, he works as an Assistant Professor at Riphah Institute of System Engineering (RISE), Riphah International University Islamabad, Pakistan. He has more than 20 publications in various journals and conferences including *IEEE Transactions*. His research interest includes human-computer interaction, artificial intelligence in healthcare, and federated learning for privacy preservation.

Bibi Saqia received her (MScS) in Computer Science from the University of Science and Technology Bannu in 2018. Her PhD is in progress at the University of Science and Technology Bannu. She has more than 13 publications in various reputed journals and conferences including *IEEE Transactions*. Her research interests include data mining, machine learning, and artificial intelligence.

Maqbool Khan received his MS degree from the Huazhong University of Science and Technology (HUST), Wuhan, and PhD degree from Nanjing University in 2011 and 2013, respectively, on fully funded China Scholarship Council (CSC) of China Scholarship. Recently, he completed his Postdoctoral Research with SCCH. He has multi-disciplinary expertise and working experience on diverse topics of big data analytics, cloud computing, predictive maintenance, explainable AI, knowledge graphs, data science, and machine learning. He has more than ten years of professional experience while working both in industry and academia. He is currently an Assistant Professor with the Pak-Austria Fachhochschule: Institute of Applied Sciences and Technology (PAF-IASST), Haripur, Pakistan, and an Adjunct Researcher with SCCH. Recently, he completed his postdoctoral research with SCCH.

1 Introduction

The rapid progress in technology and the seamless connectivity accessible through the internet, have promoted the growth of social network platforms (Sharma et al., 2018). Hence internet is saturated with an infinite number of images and videos, that is not only encourages the advancements in existing search applications and algorithms but also stimulates the development of new ones (Nada et al., 2020). However, to some extent, the increased availability of data has been linked to increased crime rates. Accordingly, the demand for detection and classification software is also growing (Doğru and Kiraz, 2018).

This study focuses on the application of convolutional neural networks (CNN) to address the problem of gender and age prediction from facial images. The problem of real-age prediction has been extensively researched by many computer vision experts (Li et al., 2019). The majority of the researchers concluded that estimating a person's age is either a classification problem or a regression problem. If we talk about age as a classification problem, we are referring to a restricted range of age groups. On the other hand, a single value is assessed for each individual if we consider age as regression (Yousaf et al., 2020). Hence, due to significant variations between different ages, it is quite difficult to determine an exact age. Amongst different feature learning methods CNN is a renowned one as it achieved noteworthy success in many applications

specifically facial expression recognition. It is a multi-layer neural network used to perform complex tasks like image analysis and classification through feature extraction which is a crucial stage due to various visual changes (Li and Deng, 2020). CNN incorporates a lot of parameters, especially as the network's depth increases and to attain optimal performance, an ample amount of training images is important.

The preceding research in this field often relied on, manually crafted facial features within controlled imaging conditions (Haines et al., 2019). Over the years, an extensive analysis including several age assessment approaches demonstrated the advancement of methods. One approach to extracting mathematical features (Li et al., 2015), i.e., measuring the size and distance of facial features, identifies face wrinkles and computing ratios among facial elements. Accordingly categorises the face into different age groups. Another approach is subspace methods projecting features into lower dimensional space, but their practical applicability in unconstrained settings was questioned (Geng et al., 2008). Some strategies involved representing face images through spatially confined facial patches however the problem of patch distribution occurred which was tackled by probabilistic Gaussian mixture models (GMM) (Han et al., 2016). For furthermore enhancement of facial features a robust descriptor was deployed instead of relying on facial features. Regardless of their significant performance, these algorithms still exhibit limitations for real-world use in unconstrained image collection scenarios. As an alternative to the GMM for capturing the distribution of facial patches, the hidden Markov model (HMM) was presented.

The study conducted in Guo et al. (2008) performed regression to estimate age and introduced the aging pattern subspace for feature extraction. They proposed a manifold learning scheme, employing support vector regression for age prediction a result of this the accuracy was quite questionable. Furthermore, the study conducted by (Zhang and Yeung, 2010) treated age estimation as a multi-task, employing a warped Gaussian process. The study conducted in Jaiswal and Valstar (2016) used an integrated approach encompassing both CNNs and bi-directional long short-term memory neural networks to extract the temporal features. Fasel methodology relies on CNNs including multi-scale feature extractors to improve the accuracy of facial expression recognition but the limitation was the extracted features were not evaluated, hence without evaluation the accuracy of extracted features is questionable (Fasel, 2002). The study conducted in Liu et al. (2015) concentrated on employing 3D CNNs to extract spatiotemporal characteristics. At the same time, temporal appearance and geometric aspects were refined by Jung et al. (2015).

Face detection in the real world is very challenging because of visual alterations in the background. One approach to tackle this problem is by training a model capable of accurately distinguishing between the face and the background in an image (Tan et al., 2017).

The rest of the paper is structured as: The first section, the introduction, covers previous research and provides an overview of the study. The second section, methodology, outlines the methods, features, algorithms, and detailed approach used in the research. The third section, dataset and experiments, presents the dataset employed, the experiments conducted, and a comparison with the state-of-the-art studies. Finally, the last section provides the conclusion, summarising the research findings.

2 Methodology

We aimed to establish a network that uses artificial intelligence to determine an individual's age and gender based on images. It involves employing current surveillance techniques to gather visual data on people who visit places like plazas and markets. Followed by age, gender, and face identification using a CNN algorithm. Furthermore, rigorous mathematical techniques should be used to assess the framework for the underlying technology.

2.1 Face detection

We used Haar cascade for the face detection feature, it is very efficient in accuracy and can run in real-time on CPU without the requirement of GPUs, changing the dimension of the facial image enables the model to detect the face more accurately and avoid the noise in the further features of age prediction and gender prediction. It makes a bounding box around the face of the user so that the background, external lighting and other features may not affect the accuracy of age and gender prediction.

2.2 Gender prediction

Gender prediction is the classification problem, it has 3 classes, male, female and transgenders. It is denoted by (0, 1, 2) the output layer consists of a softmax that is responsible for Male, Female and Transgender classes. The first output layer that uses softmax is the gender prediction output layer, It was trained on UTKFace dataset (Zhang et al., 2017) which has approximately 23706 training images with gender (male/female), and one custom dataset that contains images for transgenders which has approximately 102,647 facial images. In this module we allow the user to predict age and gender separately by using the select option so that maximum output can be achieved using this model.

2.3 Activation function for gender

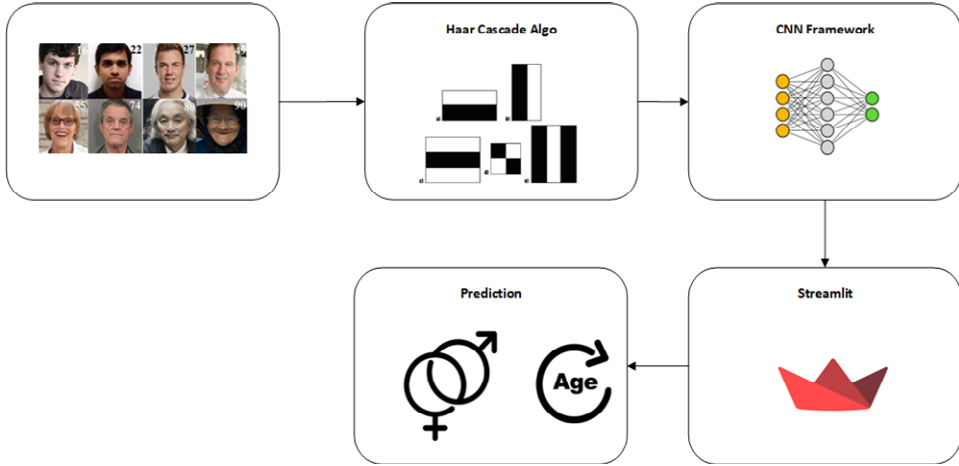
The activation function that has been used in this module of gender prediction is softmax, it is a widely used for multiclass problems that uses the probability to predict the actual class from the multi-classes, in our case the classes are (male, female and transgenders).

2.4 Age prediction

Age prediction on regression is a difficult task due to factors such as makeup, feature editing, and quality enhancement, these are some features due to which the facial features are changed in the matrix so when our model is predicted it finds difficulties with the static number so we used Adience dataset (Eidinger et al., 2014) to train the age prediction module and used another custom dataset which was created and labelled from the real world with highly accurate labels which helped the model to add more labels for age rather than only using Adience dataset, in addition it also effected the accuracy of the model in a positive way, we achieved higher accuracy by using our custom dataset with Adience dataset rather than using standalone Adience dataset. The overall model consists of three (3) convolutional layers, five (5) dense layers, two (2) dropout layers, two (2)

output layers. This module consists of less layers but higher accuracy due to higher quality of training data and validation. Figure 1 defines the steps of the proposed methodology.

Figure 1 Steps of the proposed framework (see online version for colours)



$$(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (1)$$

Equation (1) defines the softmax function used to calculate the probability distribution over diverse age classes or gender categories.

2.5 Activation function for age

Rectified linear unit (ReLU) is a popular activation function in neural networks, it is valued for its simplicity and effectiveness. ReLU introduces a non-linearity by outputting the input directly if it is positive in other cases, it produces zero (0).

Equation (2) represents the activation function for age prediction. Since we are predicting just age, we do not require any negative value because the age cannot be in negative values.

$$ReLU : (x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases} \quad (2)$$

2.6 Convolutional neural network

The CNN is a specialised type of artificial neural network that is designed for processing structured grid data such as images. CNN performs best on images as the word convolution means multiplication of matrix and images are 2d matrix. The layered architecture can vary according to the need of the problem. In this research, we propose a 12-layer CNN to predict age and gender from the facial image. In this module, we have trained separate modules for age and gender and used 2 publicly available datasets and 1

custom dataset. The proposed solution consists of 12 layers, three (3) convolutional layers, five (5) dense layers, two (2) dropout layers, and two (2) output layers (first layer for age prediction output, second layer for gender prediction output). Table 1 shows the age range of the used dataset.

Table 1 Age ranges of the used datasets

<i>Adience dataset age ranges</i>	<i>Custom dataset age ranges</i>	<i>Combined dataset age ranges</i>
(0–2)	(2–4)	(0–2), (2–4)
(4–6)	(6–8)	(4–6), (6–8)
(8–12)	(12–15)	(8–12), (12–15)
(15–20)	(20–25)	(15–20), (20–25)
(25–32)	(32–38)	(25–32), (32–38)
(38–43)	(43–48)	(38–43), (43–48)
(48–53)	(53–60)	(48–53), (53–60)
(60–100)	(65–75)	(60–100), (65–75)

2.6.1 Convolution

Equation (3) for the convolution is used for the convolution function that is performed in the model.

$$y = \sum_{c=0}^{n_c-1} \sum_{k=-p}^p x_{c,j} - k^w c, k \quad (3)$$

2.6.2 Input gradient

Equation (4) is for the input gradient which is responsible for the gradient input, this is used in the function of input gradient in the CNN, and the parameter details are mentioned in Table 2.

$$\frac{\partial L}{\partial x_{c,i}} = \sum_{k=-p}^p \frac{\partial L}{\partial_{yi+k} w_{c,k}} \quad (4)$$

2.7 Web service

We used Streamlit for the web deployment, it allows users to customise the UI in less time and with less code. The user has the option to upload the image, and that image is passed to the model after that the model applies the Haar cascade to detect the face and create a bounding box around the face then that is passed to the next module which is gender prediction, It uses CNN to extract the features and predict 0, 1, 2 (male, female, and transgenders) in parallel that image is also passed to the age prediction module which predicts the range of age and the output is displayed over the streamlit output section.

2.8 Modification in the convolutional neural network

Table 2 encapsulates the architecture blueprint of the CNN design, comprising a sequence of convolutional, pooling, dense, and activation layers. It contains each layer's parameters and output shapes. The kernel size that we have used in Conv2D is 3×3 with 64 filters and the MaxPooling2D is also 3×3 so that the feature loss can be made to a minimum for better performance and accuracy, the MaxPooling2D is 2×2 . The total number of layers in this proposed architecture is 12 with 2 output layers, one output layer is for the age prediction and the second output layer is for the gender prediction. The detailed parameters are listed below in Table 2 with 3 columns, layer type, output shape, and parameters.

3 Experiments setup

This section presents the experiments performed to evaluate the proposed solution using the benchmark techniques, details about the datasets, and the results achieved from this proposed solution.

3.1 Datasets

In the proposed study, the images employed for forecasting are resized to dimensions of 224×224 pixels. This size is selected to be harmonious with the input necessities of the fine-tuned CNN model whereas keeping a balance between computational productivity and the protection of significant facemask attributes. Resizing to 224×224 pixels certifies that the model can efficiently process the images for precise age and gender calculations. The UTKFace dataset and a custom dataset is used for the gender module that consists of 23,708 images of males and females combined, and the custom dataset contains 102,647 images of transgenders and male mixed with labels this dataset is very efficiently divided with 3 classes and does not have a problem of imbalanced classes which is the most important reason for using this dataset. Further, the images used in this dataset were normalised and outliers were removed for the better performance of the model. In addition, the Adience dataset and another customised dataset were used for the age prediction module which had the following classes mentioned in Table 3 the reason for adding a custom dataset was to normalise the imbalanced classes and add more age ranges for the better performance.

3.2 Results and discussions

We have discussed the results of the proposed model with the baseline models regarding this task. Table 3 defines the results of the proposed model than the base models. Here we have trained three (3) models and the fourth (4th) is our proposed model we have checked the accuracy of outputs, Age prediction, and gender prediction. A clear difference can be seen in terms of accuracy in outputs. Our proposed CNN is based on 12 layers in total and it has 2 output layers, one for age prediction and a second for gender prediction. Hence, we have 2 accuracies, one is for the age prediction module and the second is for the gender prediction module. The clear difference can be seen in terms of F1-score, precision, recall, and overall accuracy. The CNN model with fully connected

layers has the least accuracy in these models. After that deep CNN has performed well for age and gender prediction. Lastly, the proposed modified CNN has performed best in age and gender prediction with 12 layers.

Table 2 Proposed solution's output shape and parameters

<i>Layer type</i>	<i>Output shape</i>	<i>Parameters</i>
Input	(128, 128, 1)	0
Conv2D (64 filter, 3×3)	(126, 126, 64)	640
Batch normalisation	(126, 126, 64)	256
MaxPooling2D (3×3)	(42, 42, 64)	0
Conv2D (128 filters, 3×3)	(40, 40, 128)	73,856
MaxPooling2D (2×2)	(20, 20, 128)	0
Conv2D (256 filters, 3×3)	(18, 18, 256)	295,168
MaxPooling2D (2×2)	(9, 9, 256)	0
Flatten	(20736)	0
Dense (256 units)	(256)	5,308,672
ReLU activation	(256)	0
Dropout (40%)	(256)	0
Dense (256 units)	(256)	65,792
ReLU activation	(256)	0
Dense (128 units)	(128)	32,896
ReLU activation	(128)	0
Dropout (40%)	(128)	0
Gender output (Softmax)	(1)	138
Age output (ReLU)	(1)	129
Total parameters:		5,707,818

Table 3 Results of proposed model and other base models

<i>Input</i>	<i>Model</i>	<i>Accuracy (%) age</i>	<i>Accuracy (%) gender</i>
Facial image	Deep CNN	80.21%	86.42%
Facial image	Base CNN	49.57%	57.00%
Facial images	CNN with fully connected layers	62.38%	71.57%
Facial image	Proposed Modified CNN	91.14%	90.34%

Figure 2 indicates the accuracy of 3 different research approaches, fully connected CNN, deep CNN, and proposed CNN. Here we can see the performance of our proposed CNN is a lot better than the existing techniques with the same Epochs. This graph is used to display the accuracy of each model separately to justify the proposed modified CNN performance for the gender prediction module. Figure 3 represents the accuracy of the age prediction module, the graphs show the accuracy of 3 different research approaches, fully connected CNN, deep CNN, and proposed CNN. This graph is used to display the

accuracy of each model separately to justify the proposed modified CNN performance for the age prediction module.

Figure 2 Gender comparison with the baseline models (see online version for colours)

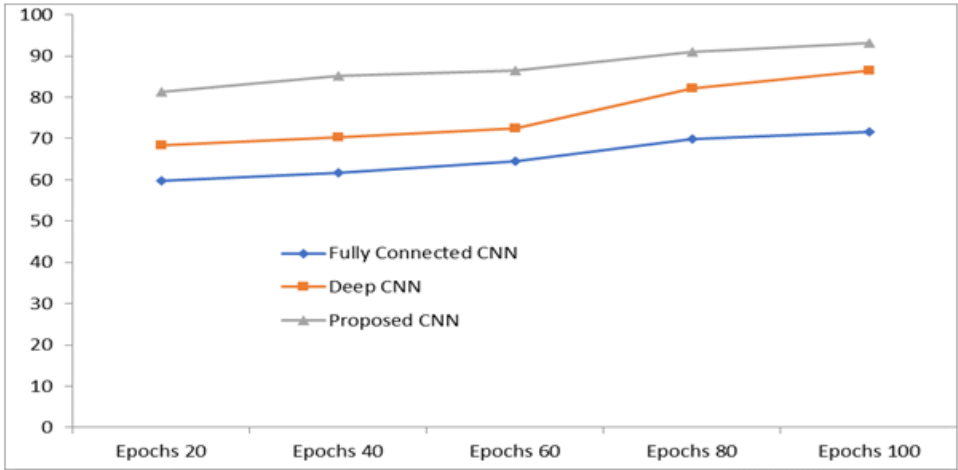


Figure 3 Age comparison with baseline models (see online version for colours)

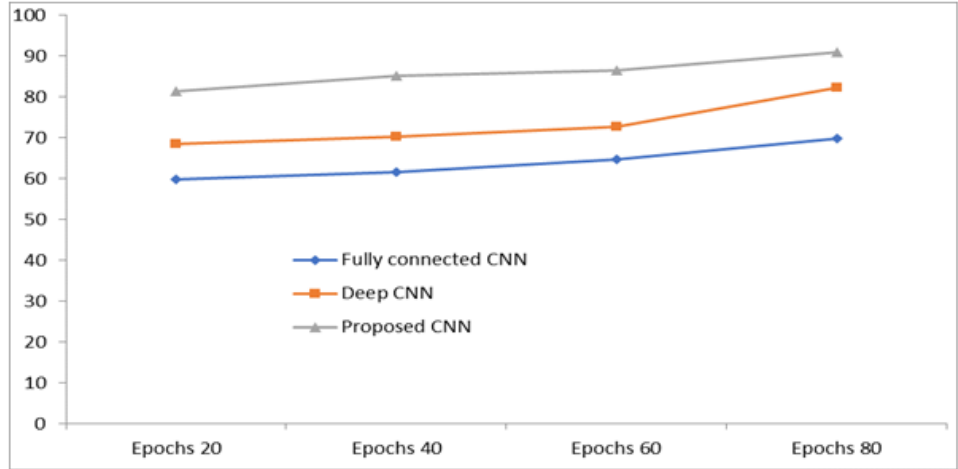


Table 4 Accuracy, F1-Score, for age and gender prediction

Model	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
Deep CNN	77.54	79.56	78.26	80.21
CNN with fully connected layers	83.27	85.47	84.70	86.42
	61.17	59.24	60.73	62.38
	69.20	69.87	68.21	71.57
	89.27	91.01	90.28	91.14%
Proposed modified CNN	90.59%	89.31%	90.40 %	90.34%

In Table 4, we have checked two models which were performing better than our proposed model and our proposed model performs better in terms of scores concerning the other available models with base architecture. The clear difference can be seen in terms of F1-score, precision, recall, and overall accuracy.

3.3 Comparison

Table 5 signifies the results of age prediction across different studies, we notice a prominent progress in age prediction accuracy over the start of the art techniques. Nada et al. (2020) attained 57.0% using CNN, while Chowdhury et al. (2020) upgraded this to 67.82% with a modified CNN. Rajababu et al. (2022) more improved performance with MTCNN, getting 72.19%. Suraj and Sapna (2023) reached 82.56% with deep CNN, V3. The suggested model in 2024 outperforms all previous methods, achieving 91.14% accuracy using a modified CNN, representing an important progression in age prediction. Table 6 presents gender prediction models that have also seen notable advancements. Nada et al. (2020) achieved 82.24% accuracy with CNN, and Chowdhury et al. (2020) attained 85.13% using a modified CNN. MTCNN by Rajababu et al. (2022) achieved 86.7%, while Suraj and Sapna (2023) gained 86.42% with deep CNN, V3. The suggested model in 2024 exceeds earlier models with 90.34% accuracy, representing the usefulness of the modified CNN for gender prediction studies. Figure 4 compares the proposed model with closely related works for age and gender prediction tasks. The graph visibly displays the incessant progression in accuracy in both cases (age and gender prediction). Gender prediction usually shows greater accuracy across all models compared to age prediction.

Figure 4 Comparison with related works for both age and gender prediction tasks (see online version for colours)

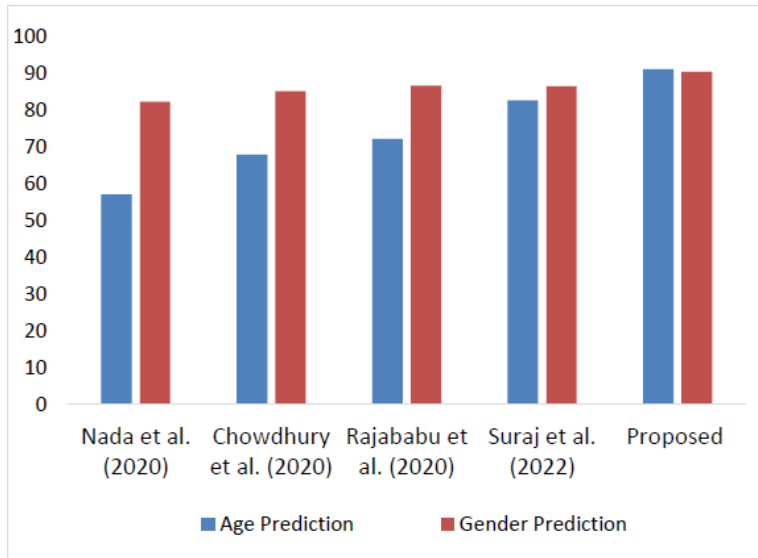


Table 5 Comparison with state-of-art-works used for age prediction

<i>Study</i>	<i>Year</i>	<i>Task</i>	<i>Framework</i>	<i>Accuracy (%)</i>
Nada et al. (2020)	2020	Age	CNN	57.0
Chowdhury et al. (2020)	2020	Age	Customised CNN	67.82
Rajababu et al. (2022)	2022	Age	MTCNN	72.19
Suraj and Sapna (2022)	2023	Age	Deep CNN, V3	82.56
Proposed	2024	Age	Modified CNN	91.14

Table 6 Comparison with closely related models used for gender prediction

<i>Study</i>	<i>Year</i>	<i>Task</i>	<i>Framework</i>	<i>Accuracy (%)</i>
Nada et al. (2020)	2020	Gender prediction	CNN	82.24
Chowdhury et al. (2020)	2020	Gender prediction	Customised CNN	85.13
Rajababu et al. (2022)	2022	Gender prediction	MTCNN	86.7
Suraj and Sapna (2022)	2023	Gender prediction	Deep CNN, V3	86.42
Proposed	2024	Gender prediction	Modified CNN	90.34

4 Conclusions

In this study, we have combined three datasets, UTKFace, Adience, and two custom dataset to train out proposed modified CNN for better performance, our age ranges consist of different age ranges present in Table 2. This model has 2 output layers, age prediction, and gender prediction, both modules are trained separately which is the reason for the better performance and accuracy. The detailed parameters are mentioned in Table 2 of our proposed modified CNN which is capable of giving an accuracy result of 91.14% for age prediction and 90.34% for gender prediction mentioned in Table 4 and Table 5. Our proposed model has fewer layers which makes it efficient and accurate. This module is less complex, requires less computational power, and consumes less time for the result. Furthermore, this module can be a very good solution when it comes to marketing purposes and targeting the audience, in addition, it can be used further for security purpose.

Acknowledgements

We honestly thank the anonymous reviewers and the editorial team for their valuable comments and direction, which significantly enhanced the quality of this paper. We also admit the support provided by Riphah Institute of System Engineering (RISE), Riphah International University Islamabad, 46000, Pakistan, which was vital for the effective completion of this research.

Declarations

There is no funding was received for this research. The authors declare that they have no conflicts of interest to report regarding the present study.

References

- Chowdhury, F.Y., Khaliluzzaman, M., Arif Raihan, K.M. and Moazzam Hossen, M. (2020) 'Pedestrian age and gender identification from far view images using convolutional neural network', *Proceedings of International Joint Conference on Computational Intelligence: IJCCI 2019*, Springer, pp.163–177.
- Doğru, İ.A. and Kiraz, Ö. (2018) 'Web-based android malicious software detection and classification system', *Applied Sciences*, Vol. 8, No. 9, p.1622.
- Eidinger, E., Enbar, R. and Hassner, T. (2014) 'Age and gender estimation of unfiltered faces', *IEEE Transactions on information Forensics and Security*, Vol. 9, No. 12, pp.2170–2179.
- Fasel, B. (2002) 'Head-pose invariant facial expression recognition using convolutional neural networks', *Proceedings Fourth IEEE International Conference on Multimodal Interfaces, IEEE*, pp.529–534.
- Geng, X., Smith-Miles, K. and Zhou, Z-H. (2008) 'Facial age estimation by nonlinear aging pattern subspace', *Proceedings of the 16th ACM international Conference on Multimedia*, pp.721–724.
- Guo, G., Fu, Y., Dyer, C.R. and Huang, T.S. (2008) 'Image-based human age estimation by manifold learning and locally adjusted robust regression', *IEEE Transactions on Image Processing*, Vol. 17, pp.1178–1188.
- Haines, N., Southward, M.W., Cheavens, J.S., Beauchaine, T. and Ahn, W-Y. (2019) 'Using computer-vision and machine learning to automate facial coding of positive and negative affect intensity', *PLoS One*, Vol. 14, No. 2, p.e0211735.
- Han, S., Meng, Z., Khan, A-S. and Tong, Y. (2016) 'Incremental boosting convolutional neural network for facial action unit recognition', *Advances in Neural Information Processing Systems*, Vol. 29, No. 1, pp.1–9.
- Jaiswal, S. and Valstar, M. (2016) 'Deep learning the dynamic appearance and shape of facial action units', *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, pp.1–8.
- Jung, H., Lee, S., Yim, J., Park, S. and Kim, J. (2015) 'Joint fine-tuning in deep neural networks for facial expression recognition', *Proceedings of the IEEE International Conference on Computer Vision*, pp.2983–2991.
- Li, H., Lin, Z., Shen, X., Brandt, J. and Hua, G. (2015) 'A convolutional neural network cascade for face detection', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.5325–5334.
- Li, S. and Deng, W. (2020) 'Deep facial expression recognition: a survey', *IEEE Transactions on Affective Computing*, Vol. 13, No. 3, pp.1195–1215.
- Li, W., Lu, J., Feng, J., Xu, C., Zhou, J. and Tian, Q. (2019) 'Bridgenet: a continuity-aware probabilistic network for age estimation', *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.1145–1154.
- Liu, M., Li, S., Shan, S., Wang, R. and Chen, X. (2015) 'Deeply learning deformable facial action parts model for dynamic expression analysis', *Computer Vision--ACCV 2014: 12th Asian Conference on Computer Vision*, Springer, Singapore, Singapore, 1–5 November 2014, Revised Selected Papers, Part IV 12, pp.143–157.

- Nada, A.A., Alajrami, E., Al-Saqqa, A.A. and Abu-Naser, S.S. (2020) 'Age and gender prediction and validation through single user images using CNN', *Int. J. Acad. Eng. Res. (IJAER)*, Vol. 4, No. 1, pp.21–24.
- Rajababu, M., Vemuri Lalitha, P.S. and Yaswanth, K. (2022) 'Multi-task CNN based age and gender recognition', *International Journal for advanced Research in Science and Technology*, Vol. 12, No. 5, pp.226–233.
- Sharma, N., Jain, V. and Mishra, A. (2018) 'An analysis of convolutional neural networks for image classification', *Procedia Computer Science*, Vol. 132, No. 1, pp.377–384.
- Suraj, S. and Sapna, V. (2022) 'Deep learning approach for human emotion-gender-age recognition', *2022 3rd International Conference on Communication, Computing and Industry 4.0 (C2I4)*, IEEE, pp.1–6.
- Tan, Z., Wan, J., Lei, Z., Zhi, R., Guo, G. and Li, S.Z. (2017) 'Efficient group-n encoding and decoding for facial age estimation', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 40, No. 11, pp.2610–2623.
- Yousaf, A., Khan, M.J., Khan, M.J., Siddiqui, A.M. and Khurshid, K. (2020) 'A robust and efficient convolutional deep learning framework for age-invariant face recognition', *Expert Systems*, Vol. 37, No. 3, p.e12503.
- Zhang, Y. and Yeung, D-Y. (2010) 'Multi-task warped gaussian process for personalized age estimation', *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, pp.2622–2629.
- Zhang, Z., Song, Y. and QI, H. (2017) 'Age progression/regression by conditional adversarial autoencoder', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.5810–5818.