Deep Learning for Age and Gender Estimation

Bharath Yalagi
Department of CSE
Ramaiah Institute of Technology
Bengaluru, India
bharathyalagi@msrit.edu

Sangeetha J
Department of CSE
Ramaiah Institute of Technology
Bengaluru, India
sangeethakirank@msrit.edu

***Abstract*: Automatic age and gender prediction via facial recognition is an important feature in modern applications, such as personalized marketing, surveillance, and security systems. This work focuses on applying CNNs in order to learn and extract automatically features from face images, doing away with manually engineered features. The system predicts age within certain ranges and accurately classifies gender, even under tough environments characterized by variable illumination, facial expressions, and occlusions. Using a diverse and labelled dataset, the model generalizes to unseen data while dealing with real-world complexities. Used Techniques like transfer learning and data augmentation enhance the robustness and accuracy of system. The suggested approach not only enhances the reliability of prediction but also shows significant promise for deployment in areas such as forensic investigations, demographic studies, and human-computer interaction. This novel approach is expected to provide a scalable and efficient framework for automated demographic analysis.**

Keywords: Convolutional Neural Networks, ResNet, VGGNet, deep learning, data augmentation

1. INTRODUCTION

Automated determination of age and gender recognition is a building application using machine learning and vision of computer techniques which has gained so much interest towards finding the ages and genders identification from images. As discussed above, applications can be related to security or even marketing tailored for individuals along with social networks, but its nature is intrinsic in complexity; there is some differentiation between almost minimal facial variation, body positions, and kinds of clothing relating to age as well as genders.

Convolutional Neural Networks are examine the most prominent way for the job. CNNs, which were developed to solve image-based applications, comprise layers such as pooling, convolutional and fully connected, designed to capture intricate patterns from input data. These layers are capable of auto-feature extraction and classification with less computational problems due to pooling, which improves prediction accuracy. Training a CNN requires large, labelled datasets to establish relationships between facial attributes and demographic traits.

In addition, transfer learning the fine-tuning of models like VGG-16 saves tremendous training period and computation energy and costs because the models exploit knowledge in pre-training. Ensemble learning using a collection of multiple CNN models boosts prediction accuracy and robustness. Techniques in data augmentation include random transformation applied to images in overcoming problems that lighting situations, faces expressions, and poses create.

Transfer learning is a further process that involves fine-tuning models like VGG-16, which reduces training time and computational demands by leveraging existing knowledge. Ensemble learning combines multiple CNN models to enhance prediction accuracy and robustness. Data augmentation techniques, such as applying random transformations to images, address challenges posed by distributing lighting instances, faces expressions, and poses.

Recent successes of CNNs have shown truly impressive performance with regards to robustness and good accuracy in age and gender detection. They also turn out to be scalable solutions to real problems if feature learning is automated and generalization for unseen data is improved. This renders them essential for modern facial analysis technologies.

The proposed model employ Convolutional Neural Networks to process facial images and extract the necessary features for classification. For accuracy, it applied the learning that is transfer learning in fine tune architectures such as ResNet and VGG16, along with augmentation of data methods that ensure adaptability to real-world variations. The Adience dataset, with labeled age and gender categories, is used for train. This provides a comprehensive dataset that permits the device to learn and generalize well. This work highlights the unique of automation in extraction of features and classification. It presents a scalable solution for demographic analysis. The belongings of the task are predicted to contribute to the field of ai and further advance deep learning for face prediction applications.

2. LITERATURE REVIEW

Other related work in this field has utilized different approaches Such as shallow learning, deep learning, and ensemble learning to predict the age and gender of an individual. The models have been trained on datasets such as which contained images of faces associated with age and gender annotations.

Irfan Rafique et al. investigated the application of DCNN's for predicting age and gender. The model achieved accuracy superior to that of conventional process, with more than 90% accuracy. The authors discussed few of the issues, such as faces variations, through data augmentation and fine-tuning, which makes robust datasets play a role in enhancing the performance of DCNNs.

Mohammed Benkaddour investigated CNN-based models in age estimation and gender classification. The study found a huge difference in the enhance of CNNs from previously old methods that demonstrate performance less than 90%. It pointed out challenges or problems, including facial occlusions and inconsistent lighting, which are recommended to be addressed through advanced pre-processing for realistic applications.

Arsala Kadri et al. discussed the ability of architectures in deep learning, including CNNs and LSTM networks, in age and prediction of gender. Their results supported the advantages of deep learning models compared to conventional methods approaches, especially concerning accuracy and adaptability. Data augmentation and model fine-tuning were suggested as strategies to cope with variations in facial features and improve robustness.

Utkarsha Kumbhar et al. emphasized the need for deep learning in demographic analysis. The CNN-based model in their research had great precision in classifying gender and age. The authors emphasized the importance of diversified datasets and proposed future work to improve the efficiency of models in fluctuating environmental circumstances, including lighting and facial expressions.

Alex Pentland et al. worked on facial recognition for smart environments, demonstrating its potential in applications such as adaptive lighting and temperature control. They addressed privacy concerns by suggesting encryption and access control to protect sensitive demographic data while maintaining the functionality of the model.

H. Takimoto et al. proposed a hierarchical approach, combining global and detailed facial characteristics for age determination. The advanced feature descriptors used in this method, such as HOG-LBP, provided great improvements in accuracy. It paved the route for more sophisticated models capable of delivering improved results, address complex facial variations.

G. Levi and T. Hassner utilized transfer learning to adjust pre prepared devices for fine tune for age and gender prediction. They use architectures like VGG-16 for extracting features and subsequently using fully connected steps for categorization. Augmenting datasets with random transformations improves the accuracy and adaptability of models.

Abu Nada et al. focused on the inquiry of CNNs for the single image-based gender and age detection. The relevance of preprocessing steps-including resizing, normalization, and histogram equalization-data enhancement was given importance in their approach, with results demonstrating the ability of CNNs to extract subtlety patterns in facial images to make precise classifications.

P. Thukral et al. developed a hierarchical regression model for age prediction using a combination of global and local facial features. Their method improved accuracy by integrating appearance and deformation models to handle diverse facial configurations. This methodology proved its value in applications requiring precise demographic analysis.

3. PROPOSED WORK

The proposed work focuses on developing a deep learning-based system for accurate gender and age recognition making use of the adience data collection. A CNN design, enhanced with transfer learning techniques like VGG-16, is added to extract the most important facial features. Data augmentation is performed to overcome problems such as lighting variations and occlusions. The system integrates preprocessing, training, and testing pipelines in order to come up with robust predictions. The ultimate goal is to achieve high accuracy and reliable performance across diverse real-world scenarios.

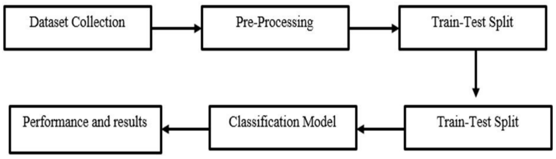


Fig. 1. Proposed model Architecture

A. Dataset

The data used in this research is the Adience dataset, comprising 26,580 labeled facial images gathered from a Flickr album. This dataset provides a wide variety of images that represent real, unaltered conditions and spans different age groups and genders. The age labels are segmented into predefined intervals, such as (0–2) and (4–6), whereas gender is either male or female. This one is designed to answer real-world questions, which means variations in lighting, occlusions of faces, and expressions, hence very representative of practical scenarios. This makes the Adience dataset a good foundation for training CNN models like VGGNet and ResNet, which is robust and accurate in demographic prediction across environments.

Fig. 2. Sample images from dataset

Fig. 3.Detected face with bouding box

B. Image Preprocessing

Preprocessing is essential in age and gender recognition using deep learning. The images are resampled to have the same size as uniformity such that it fits the CNN model requirements for the input. Subsequently, converting the image to grayscale or normalization of pixel intensity would ensure the distribution is uniform with no lighting variability. The diversity and robustness of the dataset are enhanced through techniques for data augmentation, including rotating, flipping and cropping. These preprocessing steps improve the performance and adaptability the model to real-world conditions.

C. Training and Testing :

In the training, transformed images are related to the deep learning machine CNN with their age and gender labels. During testing, unseen images are gave in the way check to verified how well the model performs on unseen data, metrics like accuracy and loss are used commonly measure how effective the system is in predicting age and gender.

D. Classification Model

i. ResNet

ResNet (Residual Networks) is an innovative architecture that deals with the concern of vanishing gradients in networks with multiple layers. This architecture allows for much deeper networks while avoiding overfitting and degradation of performance. ResNet has proven to be exceptionally efficient in many image recognition tasks, and its ability to learn intricate features makes it a very valuable model for age and gender detection in images.

ii. VGGNet

VGGNet (Visual Geometry Group at Oxford) is recognized as one of the more straightforward yet deeper convolutional neural networks, which utilize very small receptive fields essentially, 3x3 convolutions stacked on top of each other allow the VGGNet to capture complex patterns in images. Still, with its depth, VGGNet retains an uncomplicated architecture, an aspect that generally makes it both easier to implement and modify. The model achieved remarkable performance for image classification, particularly when fine-tuned for a number of particular claim such as prediction of age and gender.

iii. Convolutional Neural Networks

CNN are a class of models in deep learning designed at the outset to deal with grid-structured images data. The architecture of CNN consists of various steps, adding convolutional layers, which automatically learn spatial hierarchies from the input data, pooling layers, which lowers dimensionality, and dense layers for classification. CNN have shown effectiveness wonders in the domain of computer vision by capturing both local and global features from images tasks they are very better at, such as recognizing age or gender from faces.

4. METHODOLOGY

The proposed approach to age and gender recognition is based in the context of deep learning strategies, Convolutional Neural Networks, for the exact classification of individuals based on facial images. The methodology consists of a number of important stages, starting with data collection, pre-processing, model selection, training, and finally, evaluation.

A dataset is collected that is diverse and holds labeled photos of people from various ranges of age and gender identities. Preprocessing on the dataset normalizes the images in terms of size and color channels. This step is very crucial to enhance generalizability across different lighting conditions and facial expressions. Data augmentation is applied using rotation, flipping, and scaling to make the model more robust and Artificially pumping the quantity of the value to lower the hardship of overfitting.

The core of the model is a deep CNN designed to specifically absorb the details from images of faces. The CNN embodies multiple convolutional layers inspired by pooling layers for dimensionality reduction. These layers enable the network to learn hierarchical features such as edges, textures, and higher-level facial structures. The final the network architecture comprises fully connected layers, where the model learns to make predictions regarding both age and gender categories.

In order to make the model more accurate and efficient in training, pre-trained models like VGGNet and ResNet are employed as feature extractors. By making use of such pre-trained networks, transfer learning is facilitated, which let the for fast convergence to better performance on less training data.

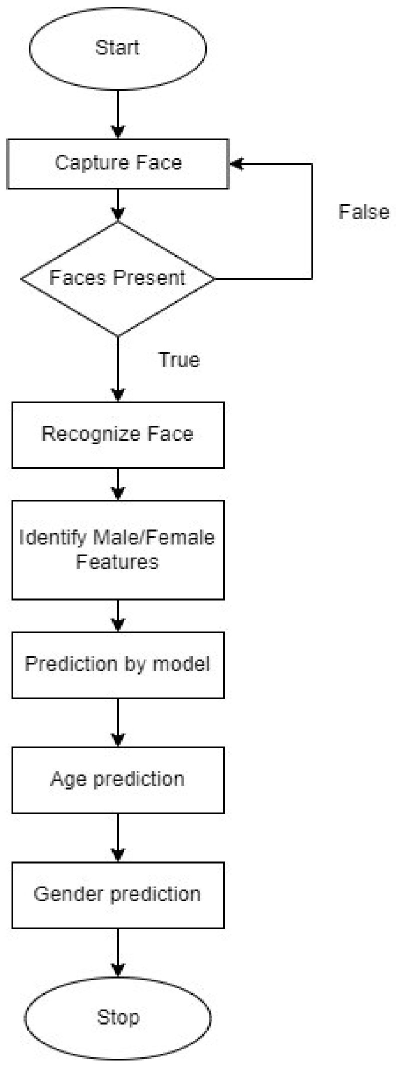


Fig.4. Workflow of the model

5. RESULTS AND DISCUSSION

The proposed model was able to successfully classify the genders with a correctness of 96.59%, and a baseline correctness of 95% is shown in Fig. 5 and gender classification loss is displayed in Fig. 6.Therefore, this degree of improvement confirms that the model is quite competent in discriminating between male and female facial images. A very high validation correctness of 86.14% was achieved that indicated generalization across unseen data. Stable convergence over 50 epochs has assured the robustness of the model in gender detection there was a small change in the accuracy of changes due to illumination, facial occlusions, among other reasons. Still, it remains highly reliable for most practical applications in real-time gender classification.

i. Gender Classification Accuracy and Performance Graphs

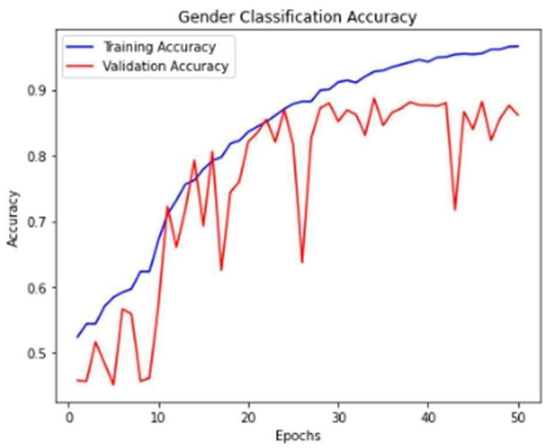


Fig.5. Gender Classification Accuracy

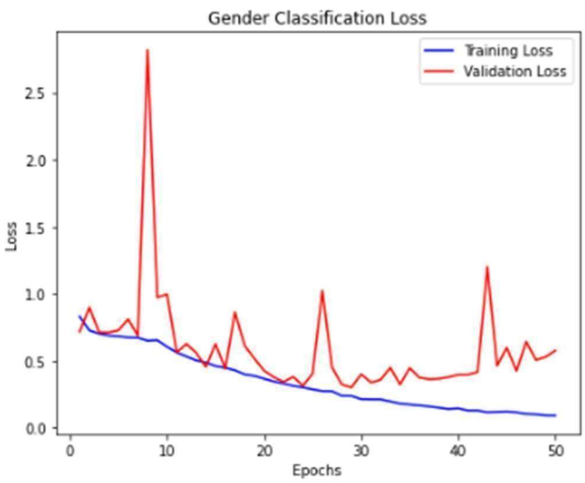


Fig.6. Gender Classification Loss

The augmentation technique introduced variation into the image. It did some rotation, scaling, and flip. Post 100 epochs, the accuracy with augmentation was shown to be about 79.32% on gender classification; the validation accuracy was recorded as 81.25% is displayed in Fig 7. And data augmentation loss of gender classification is displayed in Fig 8. Improved validation accuracy better represented adaptation in the model as it accounted for varied real world scenarios and brought down Overfitting. However, a slight decrease in training accuracy suggests that the model had a problem in learning the subtle variations introduced through augmentation. However, augmentation is a useful technique for improving robustness in gender classification.

ii. Augmented Gender Classification Accuracy and Performance Graphs

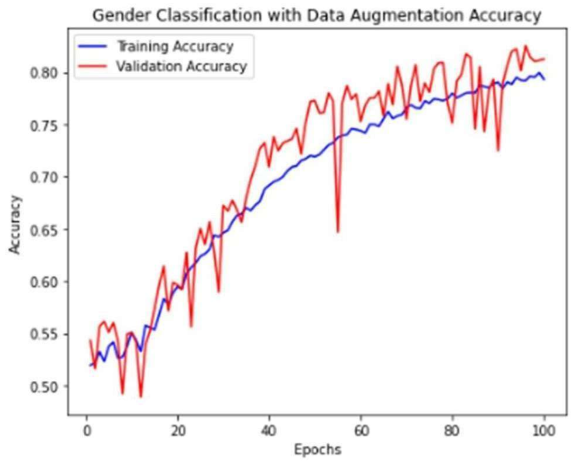


Fig.7. Data Augmentation Loss of Gender Classification

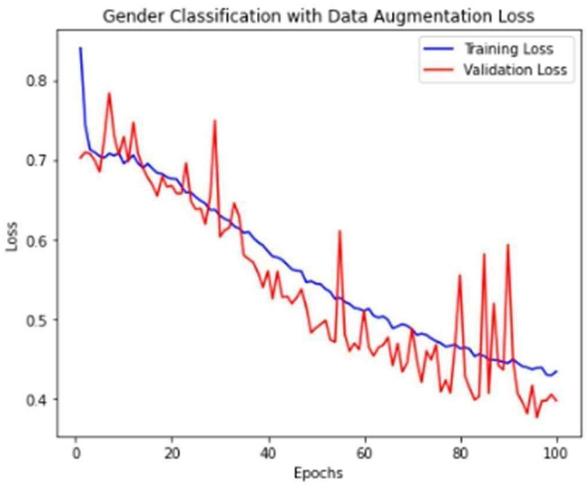


Fig.8. Data Augmentation Loss of Gender Classification

Age classification was really challenging compared with gender detection on account of the various aging processes and facial patterns. It has been noted that the accuracy the model is 62.73%. The validation accuracy after 200 epochs shows the value of 58.42% is displayed in Fig. 9 and age classification loss is shown in Fig. 10. These result in the enhancing of the aging characteristics of young age group. Since the older people share an overlapping characteristic with age, the process of classification tends to pain them. Apart from this, lighting conditions, facial expressions, and occlusions also affect the

Estimation process. However, if further developments on data balancing along with high-end augmentation techniques occur, the model can demonstrate its excellent age estimation potential.

iii. Age Classification Accuracy and Performance Graphs

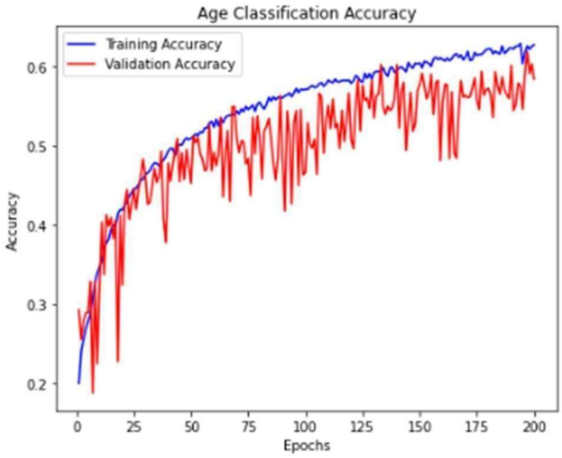


Fig.9. Age Classification Accuracy

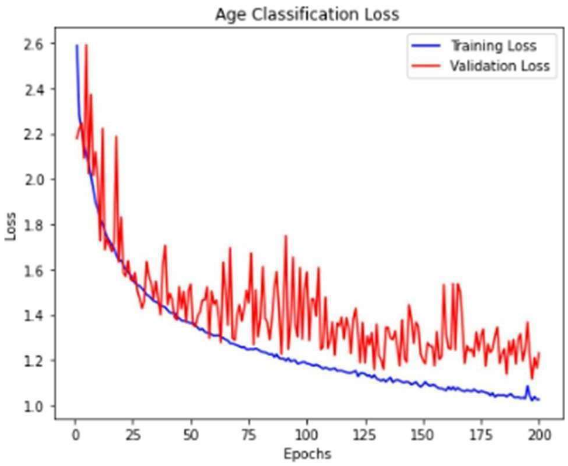


Fig.10. Age Classification Loss

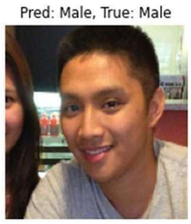
It was still not fully efficient with the images of low illumination, facial expressions, and occlusions. The predictions in such cases are usually wrong, so more enhancement is required.

Overall, the automated age and gender classification system has shown great ability and promises to be a valuable tool for diverse applications, like marketing, user interaction with computers, and biometric systems. However, additional explore is necessary to enhance its accuracy and overcome its current obstacle.

The age and gender recognition system produced in this work performed remarkably while distinguishing among male and female faces characteristics. The accuracy of the classifying gender within the system was very high based on the specific variations recognized in male and female facial attributes, which easily permitted the model to classify appropriately with minimal error.

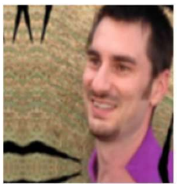
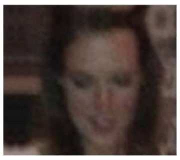
The results of the model will be shown in tabular form. The test images will be shown in Table 1 together with the relevant expected gender labels to allow for an easy comparison of input and output. Through this structured representation, the model's classification accuracy and learning efficiency are illustrated. Recall, accuracy, and precision are among the quantitative evaluation criteria that will be included to further validate the system's operation.

Table1. Testing of Gender Prediction

Testing	*Images*	*Output*
1.		
2.		

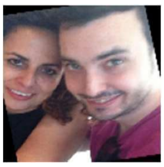
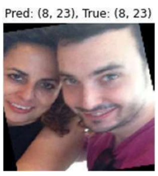
In the testing phase, the trained deep learning model is presented with an input image to assess its gender classification performance. By extracting significant facial features, the model determines whether the image is male or female. The classification results are displayed in Table 1, where the first image is the original test input and the second image is the expected gender label. The model's ability to categorize gender and its effectiveness in real-world scenarios are demonstrated by the accurate separation between the male and female categories.

Table 2. Testing of Gender Prediction for Augmented images

Testing	*Augmented Images*	*Output*
1.		
2.		

To increase the model's ability to generalize, a number of modifications were applied to the input images during the data augmentation phase. The genders of the improved images were then classified using the deep learning model that had been constructed. Table 2 displays the updated test images along with the expected gender labels. The first column displays the altered input photographs, while the second column displays the categorization outcomes. By ensuring that the model is robust to changes like as lighting, rotation, and noise, this process improves the model's performance in real-world scenarios.

Table 3. Testing of Age Prediction

Testing	*Images*	*Output*
1.		
2.		

To enhance the model's capacity to predict age, a number of preprocessing techniques were applied to the input images prior to classification. The trained deep learning model then analyzed these images and predicted the corresponding age groups. Table 3 displays the test images together with their estimated age labels. The model's predicted age classification is displayed in the second column, while the original input images are displayed in the first. This assessment shows how well the system can classify people according to age, guaranteeing that it can be used in real-world situations even when facial features, lighting, and expressions change.

PERFORMANCE SUMMARY

Model	Accuracy (%)	Validation Accuracy (%)	Epochs
Gender Detection	96.59	86.14	50
Gender Detection (Augmented)	79.32	81.25	100
Age Detection	62.73	58.42	200

The developed automated sexes and age recognition model tested on significant image datasets, whereas analysis and precision and robustness are compared largely from the open source datasets

For age recognition, the system executed the best for the young, as it was easy and clear-cut when the features due to age are more distinct and easier to predict. The system's performance has a little downward trend for the older population. The reason why the complexity occurs in the older population is due to the few of distinction between the subtle and less noticeable patterns of aging.

The system performed well in gender recognition because the facial features that distinguish males from females are usually defined and easily recognized. The system's accuracy in gender classification was very high because the features, including jawline structure, brow prominence, and facial hair in males, were defined. It should be noted, however, that performance did dip in cases with atypical or non-standard facial features, perhaps as a result of genetic, makeup, and the like variations. Overall, gender recognition accuracy still stood strong compare to the age prediction.

The system revealed a number of weaknesses in the test. It could not perform well with images whose illumination is low, whose face is changed due to different emotions, or partly occluded. Most of these usually result in false or low accuracy predictions. Improving these areas requires that better preprocessing techniques are developed and incorporated into the model, augmenting it with more robust performance to handle such variations.

The proposed model achieved better accuracy in gender classification compared with the base paper, as it reached 96.59% against the base paper's 95%. Experiments included age detection accuracy of 62.73%. Additionally, techniques for data augmentation were explored to enhance model generalization. A validation accuracy of 81.25% on the augmented gender classification model effectively demonstrated the power of augmentation against real-world variations. Although the model did well, illumination changes, facial occlusions, and age-related complexities were some of the biggest challenges. Improvement in augmentation techniques and deeper architectures may help the device become more accurate and robust in the upcoming time.

6. CONCLUSION AND FUTURE SCOPE

This paper focuses on the developments and applications of automated systems for age and gender recognition, notably in deep learning and computer vision domains. The technology has been improved significantly over the past years, with greater accuracy and reliability. However, there are still areas that need further enhancement, such as privacy concerns and ethical issues surrounding the use of such technology.The ability to predict age and gender automatically holds a great potential in most sectors. This can be of great help to businesses in targeted marketing, thereby making advertisements and campaigns more personal and appealing to certain demographics. It can also be used in healthcare to better treatment plans as it predicts a patient's age and gender for more accurate care. However, there are issues of misuse. The potential for discrimination based on age, gender, or other personal data is a real issue that must be addressed.

As the utilization of technology for automatic age and gender identification increases, developers and organizations must ensure that they adhere to high ethical standards and protect the rights of individuals by adopting privacy measures. The steps include ensuring that the technology is not misused, providing people with control over their data, and enhancing the system's accuracy. Although this technology has made significant progress, responsible implementation is essential. As used in areas that address concerns and advance precision such as marketing and the healthcare field. Real-time capabilities may yet produce a revolutionary basis for fields related to video analyses and human/computer interactions

7. REFERENCES

[1] Singla, Riya & Singh, Gurpreet. (2023). Age and Gender Detection using Deep Learning. 1-6. 10.1109/ICCUBEA58933.2023.10392246.

[2] Rafique, Insha & Hamid, Awais & Naseer, Sheraz & Asad, Muhammad & Awais, Muhammad & Yasir, Talha. (2019). Age and Gender Prediction using Deep Convolutional Neural Networks. 1-6. 10.1109/ICIC48496.2019.8966704.

[3] Benkaddour, M. K. (2021). CNN based features extraction for age estimation and gender classification. Informatica, 45(5)..

[4] Kadri, Arsala & Sharma, Kirti & Chauhan, Narendrasinh. (2019). Age and Gender Detection using Deep Learning Models. International Journal of Computer Sciences and Engineering. 7. 671-676. 10.26438/ijcse/v7i4.671676.

[5] Utkarsha Kumbhar et al Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol,
May-June - 2021, 7 (3) : 604-610

[6] Pentland, Alex & Choudhury, Tanzeem. (2000). Face recognition for smart environments. Computer. 33. 50 - 55. 10.1109/2.820039.

[7] Levi, Gil, and Tal Hassner. "Age and gender classification using convolutional neural networks." In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 34-42. 2015.

[8] Thaneeshan, Rajeetha & Thanikasalam, Kokul & Pinidiyaarachchi, Amalka. (2022). Gender and Age Estimation From Facial Images using Deep Learning. 1-6. 10.1109/ICITR57877.2022.9993277.

[9] Thukral, Pavleen & Mitra, Kaushik & Chellappa, Rama. (2012). A hierarchical approach for human age estimation. Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on. 1529-1532. 10.1109/ICASSP.2012.6288182.

[10] Gupta, R. K., Shivaprasad, M. B., & Srividhya, S. (2022). Age & Gender Detection using Convolutional Neural Network. International Journal of Engineering Research & Technology (IJERT), 11(6).

[11] Rothe, R., Timofte, R., & Van Gool, L. (2018). Deep expectation of real and apparent age from a single image without facial landmarks. International Journal of Computer Vision (IJCV), 126(2-4), 144-157.

[12] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

[13] Deng, J., Guo, J., & Zafeiriou, S. (2019). ArcFace: Additive angular margin loss for deep face recognition. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 4690-4699.

[14] Antipov, G., Baccouche, M., & Dugelay, J. L. (2017). Apparent age estimation from face images combining general and children- specialized deep learning models. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2150-2158.

[15] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks (MTCNN). IEEE Signal Processing Letters, 23(10), 1499-1503.