Multimedia Tools and Applications (2026) 85:286
https://doi.org/10.1007/s11042-026-21204-x

Global soft biometrics in surveillance: benchmark in the field & open challenges

Bilal Hassan[2] · **Hasti Soudbakhsh**[2] · **Samaneh Bahrami**[2] · **Sonjoy Ranjan Das**[3] · **Preeti Patel**[1] · **Amanullah Yasin**[4] · **Usman Khan**[4]

Received: 12 February 2025 / Revised: 26 September 2025 / Accepted: 25 November 2025
© The Author(s) 2026

Abstract

Global soft biometrics-based identification refers to the recognition of subjects using human traits such as gender, age, and ethnicity. Unlike traditional biometric methods that rely on unique physical markers such as fingerprints or iris scans, soft biometrics represents a non-intrusive, viable, and versatile approach, thus making them particularly valuable for surveillance and security applications. Despite significant advances, several issues have been associated with traditional biometrics, like maintaining accuracy, addressing algorithmic bias, and limited computational efficiency. To address those issues, this paper presents a comprehensive coverage of the current advances in Global soft biometric-based recognition as a solution, where four key contributions are made; i.e., (i) advocacy on the relevance and impact of soft biometrics in surveillance and security, (ii) development of a new and unique *CeleBImg* dataset to overcome algorithmic biases and improve diversity in soft biometric-based recognition, (iii) rigorous performance comparison of current methods in-practice for Global soft biometrics-based recognition and, (iv) identification of open challenges with potential solutions in the field within the context of surveillance and security. This paper sets a solid foundation for using Global soft biometrics in the CCTV-based surveillance and security domain, with their significance, relevance and effectiveness.

Keywords Soft biometrics · Gender · Age · Ethnicity · Recognition · Surveillance · Security

1 Introduction

The field of soft biometrics focuses on recognizing individuals based on easily noticeable and non-invasive physical characteristics (Chowdhury et al. 2023). Unlike traditional biometric systems that rely on unique identifiers such as fingerprints and iris scans, soft biometrics draws on traits including age, gender, and ethnicity (Zhao et al. 2020). Although their non-distinguishability when compared to fingerprints, these characteristics still offer critical information in terms of identification and can be meaningfully applied within different

Extended author information available on the last page of the article

Published online: 19 March 2026

contexts like tracking, identification of an individual, as well as social interactions (Huang et al. 2021). This paper has two central objectives: (i) to present a comprehensive review of global soft biometric recognition techniques—focusing on traits such as age, gender, and ethnicity within the context of surveillance and security, and (ii) to introduce and evaluate CeleBImg, a curated benchmark dataset comprising 300 publicly available celebrity images aimed at supporting fairness evaluation and demographic diversity in soft biometric research.This review focuses on the roles of age, gender, and ethnicity as a category that continues to emerge as a significant factor in the lives of people as a key trait in global soft biometric recognition systems (Li et al. 2021). With the rapid advancements in machine learning and computer vision, particularly facial recognition technologies, identification and categorization of these attributes to the commodities have been enhanced for better results. (Zhang et al. 2024). Soft biometrics allows for the creation of systems that work on unstructured environments, in which the more conventional measures can be warped because of As several studies by Choi et al. (2022) have mentioned above, human poses, lighting, and occlusion can also change over time. Age, gender, and ethnicity are important for public safety, security, and healthcare on a global scale. These attributes also find application in marketing and consumer products to integrate and run systems better (Singh et al. 2022).

These subtle features are more than the recognition of a face [1]. Though the facial features present a fundamental basis for demographic information extraction, limbs or clothes as other components of the human body are also significant identification markers (Li et al. 2021). Large datasets have been created in recent years to classify gender, age, and Ethnicity is examined through multiple modalities, and this adds complexity to recognition systems [2]. Although much advancement has been achieved, there are still issues in soft biometrics, and in the aspect of enhancing accuracy and fairness in age, gender, and ethnicity recognition (He et al. 2023). Some of these issues entail bias in datasets, model generalization across heterogeneous populations, and the requirement for robust systems that function well in real-world conditions (Zhao et al., 2020). This debate confronts the challenges listed above, assesses the present state of soft biometrics systems, and suggests possible research directions that will make the global soft biometric technologies more reliable and inclusive (Chowdhury et al., 2023). Finally, the research area of global soft biometrics, particularly age, gender, and ethnicity recognition, presents a promising alternative to conventional biometric methods (Zhang et al. 2024). While Soft biometrics has tremendous potential for numerous applications; however, developments in datasets, model calibration, and ethics are necessary to see large-scale implementation of these kinds of technologies (Li et al. 2024). Figure 1 shows the Soft Biometrics Modalities and Features Present in Human Body.

Fig. 1 Soft biometrics modalities and features present in human body (Frederik 2020)

Key contributions of this work are as follows:

- Identified key areas in surveillance and security for applying global soft biometrics i.e., age, gender, and ethnicity.
- Developed the CeleBImg dataset, a comprehensive resource for accurate and inclusive global soft biometrics-based identification.
- Performed a rigorous comparison of recent methods in practice for Global soft biometrics-based recognition.
- Identification of open challenges with potential solutions in the field within the context of surveillance and security.

2 Dataset collection and representation

In this work, we introduce CeleBImg, a novel dataset carefully curated to allow for thorough global soft biometric examination, particularly in age, gender, and ethnicity classification contexts. The dataset is made up of 300 celebrity images, which were purposefully selected to ensure a rich representation of different genders, ethnicities, and age groups. The CeleBImg dataset is composed of 300 celebrity images carefully selected from license-free and public-domain repositories. The image acquisition process ensures adherence to ethical and privacy standards, including demographic balance across age, gender, and ethnicity categories. Annotation was manually verified, and all selections avoid copyright constraints. This meticulously fit dataset should be used to assess the effectiveness and inclusiveness of transfer learning models in gender classification in different demographic groups. To standardize experimental conditions, all images from the CeleBImg database were resized to consistent dimensions for training and testing phases. During dataset development, demographic variability was deliberately included to improve real-world generalizability and reduce subject bias in biometric attributes. The following are the key contributions of CeleBImg:

- Structured gender classification: Data are clearly categorized into Male and Female classes, enabling consistent and accurate recognition modeling.
- Ethnic diversity for real-world applicability: Samples are grouped into Asian, Black, White, Mixed, and Other to reflect broad ethnic categories and improve external validity.
- Generalized age classification: Age groups are labeled as 5–15, 16–30, 31–45, 46–60, and 61 + years, supporting fair performance evaluations across age stages.

For the development conducted in this study, the CeleBImg dataset is foundational in providing high-quality, demographically diverse, and standardized input for the evaluation of soft biometric algorithms. This methodical approach establishes a robust, ethical, and inclusive basis for future research in biometric recognition systems.

2.1 Age dataset

Facial recognition usually requires the existence of large datasets covering as many phases of life as possible for accurate and generalizable model performance which, in turn, necessitates the prediction of age. As a part of systematic efforts made to age classify in CeleBImg, age attributes are represented for the dataset to ensure they possess coverage in the dataset. The Age column in CeleBImg dataset has been cut into 5 different categories for ensuring consistency and training deep learning models.

- 5–15 years: Childhood and early adolescence.
- 16–30 years: Late adolescence to early adulthood.
- 31–45 years: Middle adulthood.
- 46–60 years: Mature adulthood.
- 60 + years: Senior individuals.

This classification is consistent with prior literature indicating that facial features evolve over passively, affecting recognition efficacy [3]. We further improve the models' robustness to age-related variations by categorizing the dataset into well-defined age groups, increasing the ability to distinguish the variations by age [4]. Furthermore, as facial aging [5] is a gradual process, more accurate age classification models are achieved when narrower age ranges are used for the younger group and wider ones for older groups. The CeleBImg dataset embodies these principles and provides high quality images of faces, with variations in both the lighting, pose and facial expression [6]. Out of the CeleBImg dataset, categorization of age class enhances its general applicability in soft biometric research by establishing a commonly accepted and complete set of mechanisms for biometric age definition (Figs. 2 and 3).

2.2 Gender dataset

Gender classification is a part of global sift biometric recognition and most of the research on humans divided them into two categories i.e. Male and Female. The categorization of CeleBImg follows the above categorization so that the gender is presented with varied representational through ethnicity and age. The accuracy of the gender classification depends on many factors like the quality of the dataset, variety in the population and differences in facial features based on age, ethnic group and environmental condition [7]. As opposed to other typically used publicly available datasets like UTKFace, Adience and CelebA [8–10],

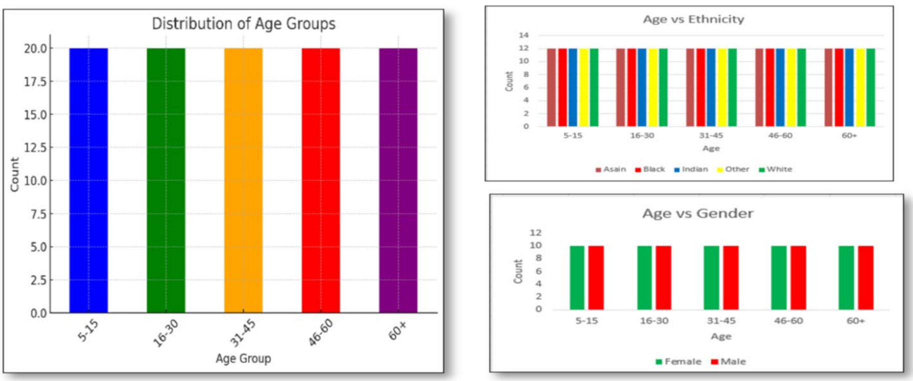

Fig. 2 Age distribution of the primary dataset

Fig. 3 Sample image from each age category of the primary dataset

CeleBImg was designed specifically to push the face recognition research frontier in the domain of soft biometrics by providing an image set with an equal gender representation. As shown by past studies, CNNs possess distinctive features in predicting valuable facial features for gender recognition with subsequent high recognition accuracy [11, 12]. However, gender uncertainty, cosmetics appearance variations caused by wearing cosmetics, as well occlusions still impact the performance of classification [1]. To tackle this, CeleBImg leverages sophisticated image preprocessing techniques (e.g. uniform resizing and pixel normalization) to increase stability during train. Besides, CeleBImg is also a useful assistant for learning a gender classification model with low bias while preserving diversity in biometrics research (Figs. 4 and 5).

2.3 Ethnicity dataset

Based on its fundamental nature of the ethnicity label in global soft biometrics, this ethnicity label provides additional information with respect to the demographic diversity in biometrics recognition system. The categorical ethnicity label system that we used can be com-

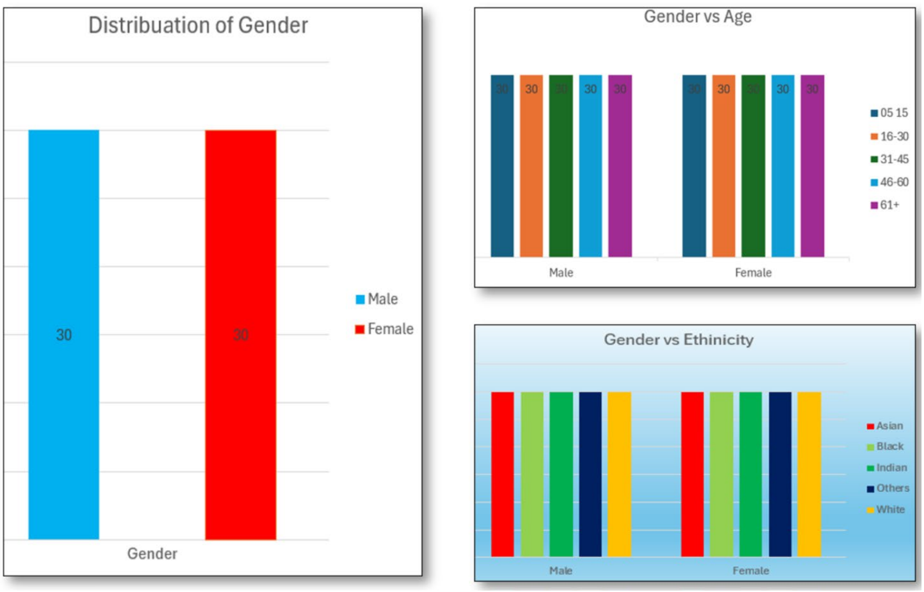

Fig. 4 Gender distribution of the primary dataset

Fig. 5 Sample image from each gender category of the primary dataset

prised of the Asian, White, Black, Mixed, and other categories for the CeleBImg dataset, that we used. It provides the means to completely represent various ethnic groups for more sensitive evaluation in biometric models. The biases in dataset have been shown to significantly impact model performance, and with a tendency to cause variability in accuracy levels by racial groups [13]. To address this existing issue, CeleBImg was designed such that it had an extremely balanced and diverse distribution that can also help reduce them in soft biometric studies. Up until now, various available dataset such as UTKFace and Fairfaced have be wildly adopted for the ethnicity classification task [11, 12, 14]. Unfortunately, it is very difficult to collect a balanced dataset of what ethnic groups. Because of the datasets with imbalances, biased performance on models prompted researchers to study fairness-based algorithms and adaptive learning strategies to maximize accuracy and inclusiveness [15]. Our CeleBImg dataset provides high-quality images which can be preprocessed with stable techniques, thus enables effective generalization of models to ethnic classification. As a valuable tool to advance the identification of soft biometrics in heterogeneous populations, it will enable fairness at the expense of reducing biases in machine learning model (Figs. 6 and 7).

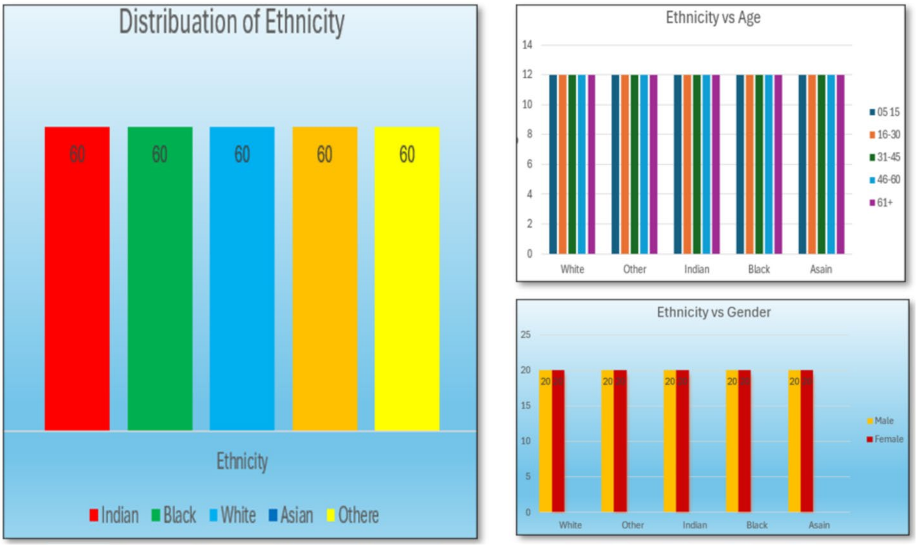

Fig. 6 Ethnicity distribution of the primary dataset

Fig. 7 Sample image from each ethnicity category of the primary dataset

2.4 Image format and quality

Images in the dataset were preprocessed rigorously to ensure a high degree of data quality. Resampling the image to JPEG and making it stand in the same dimensions of 224 × 224 pixel to make the dataset uniform. However, a rigorous examination of the images was made to discard missing values, outliers, duplicates and corrupted files in them. The final step in this process of picture preparation is that the dataset contains only consistent and good quality data for use in training and testing recognition models. However, the same consistency in picture format and size is very crucial in ensuring the success of the project due to the influence it will have on the subsequent analytical processes.

3 Significance of global soft biometric

In this present research, global soft biometrics are investigated to augment age, gender and ethnicity identification via a transfer learning approach. The efficiency of the suggested transfer learning strategy for exploiting models trained in a wide range of tasks and domains to endorse age, gender, and ethnicity classification requirements in this regard has been established. Such flexibility also allowed the use of transfer learning as an effective technique and showed its potential in enhancing model performance in other applications. Effective transfer of knowledge from a large data set into a smaller one is made possible through training the models in relevant features and patterns, which, in turn, enhances accuracy.

Today, global soft biometrics are more precious than ever before for secure and convenient verification in border control, identity issuance, smart surveillance, and e-authentication. Traditional biometrics rely on "intentional involvement of the human body" and are considered intrusive. Emerging research in non-intrusive biometrics, including soft biometrics, is paving the way for transparent systems like Smart Borders. Soft biometrics give importance to attributes like age, gender, and ethnicity, and research has investigated global estimation that is derived from images and video in various situations [16, 17].

Soft biometrics are instrumental in strengthening computer vision (CV) by mitigating its drawbacks in handling intricate and varied visual information. They augment contextual comprehension and recognition accuracy through identifiable but non-unique human characteristics. The most important facets of soft biometrics are:

- Techniques and Methods: Categorical approaches sort people into groups according to visible characteristics, comparative approaches compare features to databases, and quantitative approaches provide greater precision but struggle with uncontrolled settings (Chaudhari 2023).

- Human vs. Algorithm Performance: Research indicates that humans are better at gender identification, but algorithms are more accurate at age determination, leading to combined improvements in biometric precision (Becerra 2019).
- Reduction of False Matches: The combination of soft biometrics with conventional systems reduces errors and improves the reliability of recognition.
- Role in Security Applications: Soft biometrics facilitate transfer learning, reinforcing security and identification protocols.
- Future Potential: With improving technology, soft biometrics continue to develop, providing an increasingly sophisticated and flexible solution to identity verification.

Additionally, the integration of soft biometrics with emerging domains such as cryptanalysis for biometric security (Kaur and Kumari 2023), and the Internet of Things (IoT) for Smart Healthcare 4.0 (Sharma et al. 2024), has opened novel applications. Recent works have explored their role in Wireless Sensor Networks (WSN) for patient monitoring (Ahmed et al. 2023), and in cyber-physical systems to secure data transmission and context-aware healthcare (Roy et al. 2024). These applications highlight how soft biometrics extend beyond surveillance into secure healthcare, identity tracking, and personalized smart environments.This convergence of biometric intelligence with IoT and Big Data Analytics further enhances predictive capabilities, system interoperability, and responsiveness in both healthcare and security systems. Thus, soft biometrics stand at the intersection of privacy, intelligence, and practical security deployment across domains (Fig. 8).

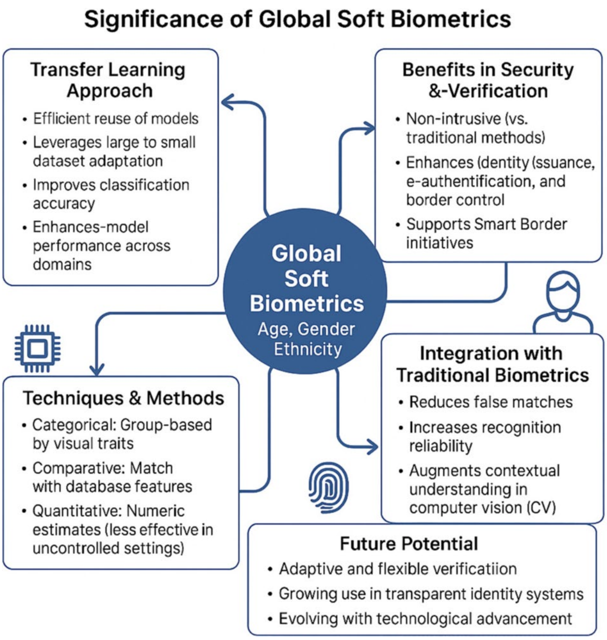

Fig. 8 Significance of global soft biometrics

4 Global soft biometrics – benchmark in the field

This study explores the utilization of soft biometrics to improve the recognition of age, gender, and ethnicity. In our research, we have found transfer learning effectively adapts models trained in different tasks, improving performance across various applications by leveraging knowledge from large datasets to smaller ones. Biometrics are crucial for secure verification in border control, identity issuance, and online authentication. Traditionally intrusive, these technologies are evolving into non-intrusive forms, with soft biometrics playing a key role in estimating age, gender, and ethnicity from images or videos for seamless, Smart Borders.

4.1 Gender recognition techniques

Recent advancements in computer vision (CV) and machine learning (ML) have improved gender identification through digital images, but challenges remain due to biases in data training and deployment. It is necessary to correct these biases to create fair and inclusive systems. Gender recognition technology is crucial in numerous fields. In internet and social media, it rendered user experiences and content suggestions more personalized, leading to higher engagement. In security and surveillance, it aids in identifying individuals, such as narrowing down search parameters in law enforcement. In medicine, it guides gender-sensitive medical treatment and research, which improves disease outcomes for conditions that have gender differences. In sales and marketing, gender demographics allow targeted marketing efforts to optimize sales and loyalty. The project aims to advance gender recognition technology to be more accurate and less biased and more equitable in terms of equality of online gender recognition.

The collection of research presented in Table 1 is new directions for enhancing the accuracy and ethicality of demographic recognition systems with a focus on bias reduction and processing efficiency. E. Kanjo, E. Younis, and C. Ang (2024) employ the AMIGOS dataset using Deep Convolutional Neural Networks (CNNs) to achieve gender recognition with a 94% rate of accuracy using physiological signals for user-interaction-based applications. Similarly, Alfonso Guarino et al. (2023) examines the utilization of gestures on mobile phones, which is translated to image data for making gender and age-group predictions. Their approach achieves as much as 94% gender identification accuracy and 99% age-group classification accuracy with CNNs and transfer learning. In their study on the importance of balance in datasets, Vítor Albiero, Kai Zhang, and Kevin W. Bowyer (2023) investigate the impact of gender balance in training data on face recognition systems. The study, which was carried out on the VGGFace2 and MS1MV2 datasets, concludes that balance in datasets does not always translate to improved performance results. Bhatta, Albiero, Bowyer, and King (2023) have looked at the contribution of hair features to the recognition system in a similar study. The authors show that changing hair attributes can eliminate gender disproportion in demographic recognition and thereby improve equity and validity. In their study, Qaswaa Khaled Abood and Farah Khiled AL-Jibory (2023) use UTK-Face dataset along with AlexNet to obtain 98 accuracy level in face recognition, proving deep learning's capability to learn different facial representations of different demographic groups effectively. That is demonstrated by Saddam Bekhet, Abdullah M. Alghamdi, and Islam Taj-Eddin (2022) on selfie photograph datasets such as Selfie, Adience, and LFW at a rate of 89% accuracy using CNNs. Just like the study done by Risky Febriawan (2022) on

Table 1 Related work for gender recognition

Ser	Year	Auth	Dataset (Sub/Image)	Features	classification/Retrieval	Performance
1	2024	E. Kanjo, E. Younis, C. Ang	AMIGOS	Physiological signals dataset	Deep Convolutional Neural Network (CNN)	Gender recognition accuracy up to 94% using the best approach on multiple gestures
2	2023	Alfonso Guarino, et al.	Mobile device gestures	Touch gestures represented as images	Convolutional Neural Networks (CNNs) with transfer learning	Gender recognition up to 94%, Age-group recognition up to 99%
3	2023	Vítor Albiero, Kai Zhang, Kevin W. Bowyer	VGGFace2, MS1MV2	Facial images	Deep CNNs with three different loss functions	Varied accuracy depending on gender balance in training data, no overall advantage for balanced training
4	2023	Aman Bhatta, Vítor Albiero, Kevin W. Bowyer, Michael C. King	MORPH, MFAD (Caucasian, African American)	Facial features, hair attributes	ArcFace, gender-balanced matcher	Gender gap largely disappears when controlling hair attributes in test data
5	2023	Qaswaa Khaled Abood, Farah Khiled AL-Jibory	UTKFace	Facial images	AlexNet	98% accuracy rate across the system's lifespan
6	2022	Saddam Bekhet, Abdullah M. Alghamdi, Islam Taj-Eddin	Selfie Dataset, Adience, LFW, FERET, NIVE, Caltech WebFaces, CAS-PEAL-R1	Selfie images	Convolutional Neural Network	89% accuracy on selfie dataset
7	2022	Risky Febriawan	Indonesian Faces	Facial features	CNN, VGG16 with Transfer Learning	High accuracy, specific rates not provided
8	2022	[16, 17]	(MMV) Pedestrian dataset	Raw pixel processing, HAAR features, LBP, texture and biologically inspired	MTCNN, ResNet, Deep-CNN, VGG-16 SVM regression, EfficientNetB3	Increasing the accuracy each time of modelling based on meter

Table 1 (continued)

Ser	Year	Auth	Dataset (Sub/Image)	Features	classification/Retrieval	Performance
9	2021	Mohammed Alghaili, Zhiyong Li, Hamdi A.R. Ali	FEI, SCIEN, AR FACES, LFW, ADIENCE	Middle face features	NN4 with inceptions	Various, up to 99.51% (e.g., FEI dataset)
10	2020	A. Greco et al.	Adience dataset, LFW dataset, MIVIA-Gender dataset, IMDB-WIKI dataset	Facial images	Optimized DCNN architecture	High accuracy with reduced computational cost
11	2020	Syed Taskeen Rahman, Asiful Arefeen, et al.	BUET facial database	Facial images	Naïve Bayes, Logistic Regression	Age classification: 76.3%, Gender classification: 86.6%

CNNs and VGG16 models using the Indonesian Faces dataset, the classification of gender practices on local characteristics among data differences can allow for better accuracy. [16, 17] use the dataset of HAAR features and MTCNN using MMV Pedestrian dataset in the domain of pedestrian detection. In their study, the traditional method and the deep learning methodology are used to enhance accuracy in pedestrian detection. In datasets such as FEI and LFW, 99.51% gender classification accuracy is achieved working in middle face features by Mohammed Alghaili, Zhiyong Li, and Hamdi A.R. Ali (2021) using NN4 and inception modules. Through using Adience, IMDB-WIKI, etc. to achieve high accuracy in an application, Greco et al. 2020 boost the efficiency of a DCNN applied to face recognition. Further, Syed Taskeen Rahman, Asiful Arefeen, and colleagues (2020) employ conventional models such as Naïve Bayes and Logistic Regression for age and gender classification with the BUET facial database. Their performance is 76.3% and 86.6% respectively, demonstrating the effectiveness of traditional machine learning methods in demographic identification.

4.1.1 Interpretation and insights

The comparison of gender recognition methods indicates deep learning-based approaches, especially those using convolutional neural networks (CNNs) such as ResNet-50 and VGG16, significantly outperform traditional handcrafted feature-based models in terms of accuracy. For instance, ResNet variants consistently show robust performance across datasets, particularly in unconstrained environments. However, many models still exhibit bias toward majority gender groups, leading to decreased accuracy for underrepresented samples. Moreover, the use of large-scale pre-trained models improves generalization but may inherit existing dataset imbalances. These observations underscore the necessity for fairness-aware training strategies and diverse datasets to reduce gender classification disparities.

4.2 Age recognition techniques

Age recognition from images is an important problem in artificial intelligence and computer vision, as we need to analyze many different images, to find the related information about the age in these images. Easily recognized age has always been a truth to bridge the gaps of available technologies as the images are of varying size and quality. In social media and digital content management, age recognition supports content moderation and age-appropriate recommendations. In security, it helps identify individuals in missing persons cases or criminal investigations. In healthcare, it aids in age-specific diagnoses and treatments, improving service quality. In retail, understanding age demographics enables targeted marketing and personalized shopping experiences. This review paper aims to develop a reliable age recognition system using advanced transfer learning techniques, enhancing accuracy and broadening the applicability of age recognition technologies.

This section summarizes recent advancements in age estimation techniques using diverse datasets. Feature extraction methods include raw pixels, appearance features, facial landmarks, and local binary patterns, with classification typically achieved through support vector machines and deep learning approaches. The outcomes of each technique are presented in the following column. Based on previous studies, particularly [18, 19] a comprehensive table summarizing age-based recognition developments from 2021 to 2024 is provided for an overview of this literature review (Table 2).

Automated face recognition systems often face challenges with age-related changes, particularly in applications such as passport renewal or locating missing people, where consistent recognition across different age stages is essential. The authors, Mittal and Patel (2023) propose a tensor subspace learning approach combined with fuzzy classification, which offers improved computational efficiency. They identify key factors influencing age-related facial recognition, including intra-class variations, individual aging patterns, and the differential impact of aging on facial features. However, this approach may not be suitable for the current project due to its complexity and reliance on a small dataset. Additionally, a review on Neural Networks for age estimation incorporates a hybrid deep learning approach with CNN-ELM for age and gender classification, where Mittal and Patel's (2023) study reveal that widening age gaps improves accuracy but reduces precision for intermediate ages, highlighting the trade-off between age gap size and recognition accuracy. In [30], a hierarchical network was developed to categorize faces into different age ranges using pre-trained 2 class CNNs of GoogleNet architecture stressing on the importance of appropriate age interval selection for accuracy. Based on their work, they give valuable insights on transfer learning and incremental CNN classification for age estimation. With the help of custom CNNs and transfer learning on pre-trained models, the proposed dual method system significantly reduces over fitting, which is a common problem in machine learning with complex image data as pointed out by Sheoran et al. (2021). VGG16, ResNet50, and SE ResNet 50 models can achieve good results for unseen samples and practical applications, showing that transfer learning is quite useful in the scenarios with low data, which is the focus point of this paper. [26] finally presents a face age gender prediction (FAGP) method based on deep convolutional neural networks (DCNN) combined with hybrid particle swarm optimization (HPSO) and support vector machines (SVM) features extraction, hence showing us another step forward in the field of age and gender classification. In essence, this method compresses dataset dimensionality and increases the accuracy of classification, performing

Table 2 Comparative analysis of age recognition studies (2021–2024)

Ser	Year	Auth	Dataset (Sub/Image)	Features	classification/Retrieval	Performance
1	2024	[20]	UTKFace, CASIA	(CNNs) for Facial features,	VGG16, ResNet50, Mobile Net, and VGG19	The VGG16 model on the UTKFace dataset achieved an MAE of 1.76, and the Mobile Net model on the Casia Africa Face Dataset had an MAE of 1.10, indicating high precision in age predictions.
2	2023	[21]	FG-NET and MORPH II	Landmarks detected with MTCNN, ResNet-18 model on ImageNet	MV, Contrastive learning from triplets of face images.	Accuracy percentages are not provided
3	2023	(Mittal & Patel 2023)	FG-NET, AGEDB, and MORPH-II	LBP, PCA, LDA	Fuzzy synthetic and Tensor Subspace	99.15% for FG-NET, 99.20% for AGEDB, 99.8% for MORPH II
4	2023	[22]	UTKFace	DNN, convolutional layers (Conv2D, pooling layers (MaxPooling2D)	Deep CNN	MAE (6.5) and 90% accuracy for age recognition
5	2023	(Logronio et al. 2023)	Own dataset with 640 samples which 320 is unique	Facial features, OpenCV	Keras with CNN	84.38%
6	2023	[23]	Own dataset with 3561 image sample	Extracting various semantic features	MTCNN, DELWO	For Face: 98.50% For Age: 82%
7	2023	(Vankayalapati et al. 2023)	Audience	Haar Cascade	CNN, Coffe model, OpenCV	85%
8	2023	[24]	UTKFace	Facial attributes, pixel-level	Keras library	For age: 82.5%

Table 2 (continued)

Ser	Year	Auth	Dataset (Sub/Image)	Features	classification/Retrieval	Performance
9	2023	[25]	Obtained from Kaggle.com and used 5,000 face images	Facial feature	CNN, EfficientNetV2B1	CNN method offers better results in age estimation process and accuracy compared to the transfer learning method
10	2023	(Sathyavathi and Baskaran, 2023)	UTKFace, IMDB-WIKI, CACD, FG-NET	Gabor filters, PCA, texture and orientation features	CNN DNN LSTM proposed work "SLSTM-DNN", Specially for age from human face image	81.32 82.34 88.12 91.45
11	2023	(Manlises et al. 2023)	FG-Net, Adience	Facial features	Keras and DCNN models	(MAE): DCNN: 16.86% Keras: 6.61% Concluded: Keras model was more effective in age classification
12	2022	[16, 17]	(MMV) Pedestrian dataset	Raw pixel processing, HAAR features, LBP, texture and biologically inspired	MTCNN, ResNet, Deep-CNN, VGG-16 SVM regression, EfficientNetB3	Increasing the accuracy each time of modelling based on meter
13	2022	[26]	Adience UTKface	GO, PSO, ACO, HPSO	VGG16 VGG19 Inception V3 Proposed age classification module	87.93 89.48 93.78 97.03
14	2022	[27]	MMU, CASIA, UBIRIS, and random datasets	Raw pixel, LBP, HAAR	ResNet 50, VGGNet, UNet,	89.6, 95.3, 93.1, 95.62, and 93.40
15	2022	(Sharma et al. 2022	UTKFace	Facial feature	CNN utilizing SoftMax	For age: 94.01%
16	2022	(Vidyarthi et al. 2022)	UTKFace	MTL, Shared low-level features	EfficientNetV2B1	MAE: 0.063 90.31% accuracy

Table 2 (continued)

Ser	Year	Auth	Dataset (Sub/Image)	Features	classification/Retrieval	Performance
17	2022	[28]	UTKFace	The custom CNN model consisted of an input layer, five convolutional layers, three pooling layers, a flattened and dropout layer, one fully connected layer, and an output layer. MobileNet utilized depth-wise separable convolutions	MobileNet, Custom CNN model	MAE for age: 5.3381, The custom model gave an MAE for age: 6.7615
18	2022	[29]	UTKFace	LBP, HOG, fscmrmr	SVM	95.69% accuracy
19	2021	(Yaman et al. 2021)	FERET (age)	Feature fusion convolutional layers 3 fully connected (FC) layers	VGG-16 ResNet-50	63.73% 62.37%
20	2021	[30].	FGNET and MORPH	AAMs	Two-class CNNs (Google-nets) Res-Net18 Alex-Net VGG16	89% 88% 87% 86%
21	2021	(Sheoran et al. 2021)	UTKFace	LBP, PCA, LDA	Best Custom CNN VGG_f_age ResNet50_f_age SENet50_f_age	(MAE) 5.67 4.86 4.65 4.58
22	2021	(Muneer et al. 2021)	CACD, MORPH	Conditional Generative Adversarial Networks	TensorFlow, Keras	It did not mention the number, but the paper insists that GAN can affect the result.
23	2021	[31]	FG-NET, AFAD, Wikipedia DB, UTK-Face, AdienceDB	Face structure and texture	GRA_Net	(MAE) For: FG-NET: 3.23 AFAD: 3.10 Wikipedia Age: 5.45 UTKFace: 1.07 AdienceDB: 10.57

Table 2 (continued)

Ser	Year	Auth	Dataset (Sub/Image)	Features	classification/Retrieval	Performance
24	2021	(Rajpoot et al. 2021)	FG-NET, WEAFD, UTKFace	Geometric features, textural features, LDA	CNN, ResNet, Wide ResNet	92.5%
25	2021	(Vasavi et al. 2021)	APPA-REAL and own dataset with 40 samples	dlib detector, SART	Wide ResNet model	MAE for the APPA Real dataset was 1.61, and for the own dataset, it was 1.54

much better than state of the art in precision and recall. The results of [27] were not impressive (95.6% accuracy) versus pre-trained models: 1) ResNet-50 and VGGNet. Evaluated on various optimizers (e.g., ADAM) for the price, their model had excellent performance on benchmark datasets, indicating its practicality in the real world. However, with security and passport verification as age sensitive applications, the study brings to the forefront the need to not use the same data splits to prevent data leakage and ensure the validity of the performance in any case. This project is recommended to be optimized by ADAM.

In contrast to the previous author, [21] use cosine similarity and triplet margin losses with contrastive learning to generate age discriminating features for MORPH II and FG-NET datasets. This review paper analyzes their approach, which achieves state of the art performance, through Grad-CAM and sees that their model concentrates on different facial features depending on the age group, and that the MV approach treats the face as a whole. Therefore, this provides his virtues both as an engineer and as a scientist. The approach goes about as follows: face determination utilizing MTCNN, image standardization, feature extraction utilizing an instructed ResNet-18, also plan inclusion during preparing utilizing the Adam optimizer. Such an answer is relevant to the project. [32] review CNN architectures such as AlexNet, VGGNet, GoogLeNet, ResNet, for age estimation and, through comparison, find the strengths and limitations of each. In addition, they also present aging databases and strategies, which allows a better selection of the best CNN architecture for the age detection in this project. In [22], the authors present a novel approach for age and gender classification via DCNN fine-tuning on the UTKFace dataset. From traditional feature-based methods this approach improves accuracy and scalability. With substantial applications in social media and security, there is need for precise age and gender classification, and this study puts much emphasis on the same. The CNN used to predict age is fine tuned to improve accuracy and computation time. By changing the network size, the number of the filters and depth, they optimize the network to work for real cases like video surveillance and checking age restriction in apps. It is shown that compared to the conventional methods, while the model's accuracy has a significant jump, it also solves the problem by maintaining that level of accuracy without modifying the technology stack. Likewise, Rajpoot et al. (2021) target efficient CNN for real age prediction, crucial for utilization cases such as social media moderation and security. By exploring facial features such as eye spacing, nose width, skin texture, their model runs quickly while still achieving high accuracy. To make our model more accurate across different groups demographics, it is trained on UTKFace, FG-NET and Adience.

Logronio et al. (2023) integrate CNN with Keras for the age range classification. Using preprocessing steps like resizing, gray scale conversion, cropping to extract the most features from a facial image of Filipino citizens across various age groups, the model describes. In real time it tested the trained model using the Raspberry Pi Camera and 70% accuracy threshold. A practical solution to real time age detection from facial recognition is provided by this combination of software and hardware. [23] conducted another study, where it collected huge number of images from the King Abdullah bin Abdulaziz AL Saud biography documentation center which yielded trillions of images. With the data extracted using a face detection algorithm, they first pre-processed the data with typical data augmentation techniques such as augmentation to reduce sensitivity to rotation, cropping, and noise. For face detection, they used MediaPipe and MTCNN and MTCNN gave the best results. We used the DELWO system for face recognition using pre-trained models like VGG16, VGG19 etc. Furthermore, the system was also able to predict King Abdullah's age in different images with a 98.5% accuracy in face recognition and 82% accuracy in age estimation, considering the image quality. There is a further research requirement to improve results using higher quality data. In this paper, Muneer et al. (2021) use Generative Adversarial Networks (GANs) to preserve the identity of a subject while calculating age progression for face images. In this case, they used the Cross_Age_Celebrity_Dataset (CACD) with 17,800 images and preprocessed using strict preprocessing including resizing, cropping, and normalizing. The architecture of their GAN contained a generator generating aged images and a discriminator that distinguished real and fake images to refine the results in an adversarial manner. The architecture of GAN system consists of generator and discriminator which are designed as convolutional neural networks. By using transposed convolutional layers, the generator encodes input aging images with stable facial features such as identity as the target. The discriminator, cross entropy classified, is trained to classify real images vs. generated images so that it can help improve the generators capacity to make realistic aged images. Through adversarial learning, the study presents an identity preserved loss function to preserve identities during the aging synthesis. The training environment was Python, TensorFlow, Keras, etc. open-source libraries and the GAN worked well for generating very realistic aged facial images. Based on the idea of Gated Residual Attention Network (GRA_Net) for age recognition, [31] propose. On the way to training images standardization, model processes in total five age-datasets (FG-Net, Wikipedia, AFAD, UTKFace and Adience DB), all of which require images standardization. The network gating in GRA_Net enhances learning efficiency to improve feature learning and decreases the number of units in hidden layers, while also increasing skip over layers for better transmission of information. At inference and backpropagation stages, it uses dual branch attention to refine features and improve the age prediction accuracy. The spatial and channel attention mechanisms can discriminate between aging groups by considering the key features. The model also calculates Mean Absolute Error (MAE) as the basis for evaluating performance. To predict age and gender from camera images or video streams, Vankayalapati et al. (2023) employ CNN in Keras. Image pre-processing, feature enhancement and removal of noise is first done. Then feature extraction is performed with highlighting of main facial traits using edge detection techniques and Haar like features with AdaBoost classifiers to enhance face analysis.

It then addresses the challenges faced in real time facial recognition such as variation in the brightness, the accuracy of the background and posing changes which affect the image quality. Despite that, the CNN model has performed up to 85% accuracy in predicting the age and gender using pre-defined classification Caffe models. This development utilizes deep learning, computer vision and image processing techniques, which make them applicable in real life applications such as airport security, law enforcement, and age restricted access control. This represents a major step forward in facial feature analysis and machine learning applications. However, [24] propose methodology with collection of data, feature engineering and segmentation. Convolutional and max-pooling layers are used to extract features from the images, and they are preprocessed to be classified using CNN, making arrays of them. Keras facilitates the modelling process with various activation functions like ReLU and SoftMax. On the other hand, the CNN models achieved 82.5%, 89.45% and 86.69% for age prediction, gender identification and ethnicity estimation respectively. The study notes that the illumination and resolution can affect the image quality, however their wide approach of training and real time testing makes it a very important step towards applying deep learning to face analysis for practical use in security and biometric systems. The authors of Sharma et al. (2022) improve an age and gender prediction CNN on the UTKFace dataset that covers ages from 0 to 116 years. Age and gender classification is done through convolutional layers with different filter sizes, ReLU activations and Max pool layers to extract features from the first feature section. Age prediction reaches 94.01% and gender prediction 99.86%. The model is demonstrated to be robust to real world use on cross dataset IMDB-WIKI, CACD, and FG-NET and to address image challenges like occlusions, different poses, and makeup. Thus, [20] tackles the age estimation from facial images using four CNN architectures such as: VGG16, ResNet50, MobileNet, and Xception. They design a very thorough data preprocessing pipeline to include face detection, alignment, normalization and data augmentation. To improve transfer performance, transfer learning resorts to reusing features learned in a large dataset such as ImageNet for the age estimation task. Gradient descent, backpropagation, Adam optimizer are used in the optimizing process to improve learning. On the Casia Africa Face Dataset, the MobileNet model attained a very reliable Mean Absolute Error (MAE) of 1.10.

Age and gender classification with multimodal biometrics, using ear and profile face images, is done in multimodal ear and profile CNNs by Yaman et al. (2021). Datasets such as FERET, UND-F, UND-J2, etc. on which fusion methods are applied and models like VGG-16 or ResNet-50 are used. Fine tuning at two stages enables better adaptation to new domains and higher capacity. The classification accuracy of the study is high and superior to that of unimodal approaches. Spatial fusion in VGG-16 and feature fusion in ResNet-50 achieve the best performance when it comes to age classification. For age recognition, ResNet-50 seemed to be quite promising in accuracy. [33] implement another study that uses transfer learning through Keras; pre trained networks such as VGG-16, VGG-19, and ResNet-50 parameters are adapted for age and gender classification. Lastly, their results show that by using deeper networks such as VGG 19, we can get an accuracy up to 95% for ResNet 50. We suggest this approach to study the importance of architecture's adaptation to certain tasks and we argue for more work on the study of transfer learning in real world applications considering ethical issues. Beyond that, Sathyavathi & Baskaran (2023) presented an age recognition DCNN-CS model. First, we preprocess facial image, which includes grayscale conversion, histogram equalization, and noise removal, etc. Face edges were detected by Gabor filters, and the dimensionality was reduced using PCA. The Cuckoo Search algorithm is used to improve feature learning and optimize computation by DCNN-CS model. Results show that the model outperforms traditional models on both accuracy (91.45 vs. CNN (81.32%), DNN (82.34%), LSTM (88.12%), time execution (0.001s vs. CNN (3.501s), DNN (3.501s), LSTM (1.139s). Two different methods for facial age recognition are present in Manlises et al. (2023) with the first being based on Keras and the second using traditional CNNs. Training and evaluation are carried out on FG-Net and Audience datasets to make sure the comparison hypothesis is fair. The same sentence is then followed by image preprocessing (resize, normalize) and then feature extraction (convolutional layers). With a mean absolute error (MAE) of 6.61%, a Keras model on top of TensorFlow was built and with an MAE of 16.86%, the CNN model was achieved. This demonstrates that Keras surpassed the other model by superior accuracy and time efficiency. They concluded that expand research to other datasets or frameworks to test this finding to improve age classification performance.

A robust age recognition method is proposed in [29] including preprocessing, feature extraction and classification with Support Vector Machines (SVM). First, facial images are normalized to have uniform size and intensity; secondly, features are extracted using LBP and HOG due to the robustness of lighting changes. Minimum Redundancy Maximum Relevance, a feature selection optimalization is applied on feature relevance. To test the model, we have used the UTKFace dataset, and feature selection + SVM achieves a high accuracy of 95.69%. In [28], a custom designed CNN model is compared with a MobileNet based pre-trained model using the transfer learning for age prediction. Both models are applied to the UTKFace dataset, which has 23,708 images labeled for age, gender, and ethnicity. MobileNet is one among 13 layers for the custom CNN architecture with feature extraction, depth-wise separable convolutions (optimal for mobile applications).

Lastly, the custom CNN attained a mean absolute error (MAE) of 5.3381 while the MobileNet marginally came up short with a MAE of 6.7615. This work presents how such transfer learning using pre-trained models like MobileNet significantly decreases the computational overhead as well as increases the precision of age prediction. Like Vidyarthi et al. (2022), Vidyarthi et al. (2022) also used the UTKFace dataset and proposed a model based on pre trained CNN architecture such as EfficientNetV2B1 for joint age and gender classification. Finally, the model uses transfer learning by freezing the layers of pre trained model and training new trainable layers on UTKFace dataset. The model achieved a gender classification accuracy of 90.31% and mean absolute error of 0.063 for age estimation using global average pooling, dropout and fully connected layers. It proves that the combination of both multitask learning and transfer learning could increase the accuracy of age and gender recognition models. A modified Wide ResNet architecture, a method for robust age recognition, is proposed in Vasavi et al. (2021). The preprocessing of CCTV video frames starts with converting it to grayscale, resizing and data augmentation. To overcome the vanishing gradient problem, it adds layers which extract invariant features and reconstruct images. The method varied up to 5% on the accuracy of the model compared to other methods where the model classified age into 101 classes (0 to 100 years). As age prediction in surveillance applications is emphasized in the study, it states that Wide ResNet and advanced preprocessing techniques show the greatest potential.

[34] compare two methods of age estimation: traditional CNNs vs. transfer learning with EfficientNetV2B1 on Kaggle dataset which has nearly 200,000 images (reduced to 5000 in the experiments). Feature extraction is done after the preprocessing steps such as resizing and face detection. Using convolutional and pooling layers, the CNN model is built from scratch and learns age related features. An EfficientNetV2B1 transfer learning model pre-trained on large datasets is fine tuned for age estimation. RMSE and R squared are used to evaluate the performance of traditional CNNs and it is observed that although the performance is slightly better, transfer learning helps in faster training and reduction in overfitting. It shows the tradeoffs between the two methods from an accurate efficiency point of view.

4.2.1 Interpretation and insights

Age recognition performance shows considerable variation based on architecture and dataset complexity. CNN-based models again outperform traditional approaches, with models like MobileNetV2 offering a trade-off between computational efficiency and recognition performance. However, misclassification rates tend to rise in age boundary groups, particularly between adolescence and early adulthood (e.g., 15–25 years). This indicates challenges in modeling the gradual facial changes associated with age. The results also highlight that fine-tuning pre-trained networks on age-specific datasets yields better generalization. Addressing the age ambiguity challenge may require hybrid architectures that incorporate additional contextual or temporal information.

4.3 Ethnicity recognition techniques

Ethnicity recognition is essential in personal identification, behavioral traits, social group affiliations, and cultural connections. Despite challenges like feature complexity, class imbalance, and limited datasets [35], its applications are significant across various fields. Its use in healthcare helps us understand disease prevalence and adjust treatment to increase patient care. It also helps with the surveillance system for better monitoring and threat detection [36] in security. In business, ethnicity recognition enhances targeted advertising, allowing companies to tailor strategies to specific demographics. It also improves user interaction in human-computer interfaces by fostering inclusivity. In fashion, it plays a role in visagism, guiding the selection of accessories and makeup that enhance individual features, ultimately improving customer satisfaction. In conclusion, ethnicity recognition through computational methods provides advantages in healthcare, security, marketing, and more, enhancing personalization, security, and user experience.

4.3.1 A comparison of ethnicity recognition technique

In this session, a comprehensive view of the related previous works since the year 2020 and onward with the key issues and successes are going to be discussed based on deep learning and transfer learning methods, facial recognition, and ethnicity recognition. The surveys have addressed the significance of ethnicity recognition, its historical context, different classifications, and features within deep learning methods (Table 3).

For developing a model to assess facial attraction, [43] developed a facial attraction model using transfer learning with CNN, Xception, and an attention mechanism on the

Table 3 Summary of ethnicity reviews from 2020 to 2024

Ser	Year	Auth	Dataset	Features	Classi/Retri	Performance
1	2024	[37]	Unique dataset	Facial features	EfficientNetB4, ResNet-50, SqueezeNet, VGG16, and MobileNetV2	Acc: 96.73%, 94.91%, 93.39%, 92.48%, and 90.32%
2	2024	[38]	Unique dataset	Not mentioned	A deep learning technique called Multi-Axis Vision Transformer	Acc: 77.2%
3	2023	[39]	RFW dataset	Racial Gradation	ResNet50	Acc: 84.31% (African)−76.92%(Asian)−79.38%(Caucasian)-79.47%(Indian)
4	2023	[40]	UTKFace and Fair Face datasets	MLBP, HOG, Color histogram, and SURF-based	SVM One Versus All classifier	Acc: 89.14%, 82.19%
5	2023	[41]	MORPH and FERET datasets	Face's central region	Convolutional Neural Network (CNNs)	Acc: 86%
6	2023	[42]	Unique dataset	Multi-task Cascaded Convolutional Network (MTCNN)	Convolutional Neural Network (CNN)	Acc: 87.3%, 56%, and 56% for the ethnicities Hausa, Igbo, and Yoruba
7	2023	[43]	SCUT-FBP5500 and Chang Gung Memorial Hospital Taiwan datasets	Sharpening filter and Smoothing filter	CNN (Convolutional Neural Network) and Exception	MAPE: 27.03 and 24.74
8	2023	[44]	Hybrid dataset	Pre-trained ResNet model	DCNN (Deep Convolutional Neural Network)	Acc: 89.25%
9	2023	[45]	Unique dataset	Google Teachable Machine	MobileNetV2 and DenseNet169	Acc: 100%
10	2023	[46]	P-DESTRE dataset	Individual's skin tones	Not provided	Acc: 98%
11	2023	[47]	Cupid dataset	Multi-layers feature fusion	Self-supervised network for Fine-grained Ethnicity Classification (SS-FEC)	Acc: 72.52%
12	2022	[16, 17]	MMV Pedestrian dataset	Intrusive features	One Detect, EfficientNetB3 architecture, SVM and VGG16	Not provided
13	2022	[48]	LSAFBD dataset	Efficient Net	E-BLS ER-BLS	Acc: 73.13% 74.69%
14	2022	[49]	VMER dataset UTKFace dataset.	HHODTLF-FER model	BiLSTM	Acc: 99%
15	2022	[50]	Unique dataset	Not provided	DCNN model using FPGAs	Acc: 96.9%

Table 3 (continued)

Ser	Year	Auth	Dataset	Features	Classi/Retri	Performance
16	2022	[28]	UTKFace dataset	Raw pixel	Pre-trained MobileNet model	Acc: 80.22%
17	2022	[35]	UTK Image dataset	Histogram equalization neural network layers	Ensemble Convolutional Neural network	Acc: 78.9%
18	2022	[36]	UTKFace and FairFace datasets	Middle part of the face with CNN, HOG and SVM	CNNs	Acc: 80.34%
19	2021	(Al-Humaidana et al. 2021)	Unique dataset	Histogram of Oriented Gradients and Linear SVM method	Deep clustering	Acc: 59%
20	2021	(Khan et al. 2021)	FERET, CAS-PEAL, VN-Faces, VMER datasets	Facial feature discovery Framework and Probability maps (PMs)	Deep convolutional neural network (DCNN)	Acc: 100%, 99.2%, 92%, 93.2%
21	2021	[51]	CJK and RoC datasets	Denoising convolutional autoencoders	Ensemble of Convolutional Autoencoders (E-CAE)	Acc: 80.69% and 61.81%
22	2021	[52]	Unique dataset	Off the shelf face detector	CNNs the Inception Resnet-v2 network	Acc: 96.64%
23	2020	(Brömme et al. 2020)	Unique dataset	RGB pixel	VGGNET architecture	Acc: 96.71%
24	2020	(Samir Brahim et al. 2020)	CUHK and KFDB datasets	DCNN VGG 16 and Gabor filter	SVM	Acc: 92.5%
25	2020	[53]	Unique dataset	Clustering	Deep Convolution network	Acc: 80.5%
26	2020	[54]	UTKFacce	Regression to classification and data rescaling	Customized CNN	Acc: 78.88%
27	2020	[55]	Unique dataset	RGB images	Customized CNN	Acc: 96.9%
28	2020	[56]	VMER dataset	OpenCV based on ResNet-10 detector	VGG-Face	Acc: 94.1%

SCUT-FBP5500 dataset, achieving 0.50 RMSE and an 18.5% MAPE error. [45] introduced MobileNetV2 and DenseNet169 models with a new dataset, achieving 100% accuracy using Google Teachable Machine. [48] proposed ER-BLS, combining EfficientNets in transfer learning with Backpropagation Least Squares (BLS), achieving 74.69% accuracy, though this result was not high enough. Dhanaraj and Srid [44] used five pre-trained DCNN models, including EfficientNetB7 and ResNet50V2, with a hybrid dataset, improving accuracy to 89.25%. [37] implemented transfer learning with multiple pre-trained models, yielding accuracy ranging from 90.32% to 96.73%. [16, 17] proposed OneDetect, a federated learning system for identity verification, incorporating global soft biometrics like gender, age, and ethnicity, using the MMV Pedestrian dataset. For ethnicity recognition, [38] achieved

77.2% on ethnicity recognition task using the Multi-Axis Vision Transformer. Using skin tones as features, [46] achieved 98% accuracy, but Brömme et al. (2020) observed possible biases. Finally, in [42], face and landmarks detection were made in sequence and with precision with the help of the MTCNN method using CNN models such as ResNet 50 and EfficientNet to forecast an African sub-ethnic category from a Nigerian dataset.

In [36] they proposed ethnicity classification method based on middle facial region which achieved 80.34% for five class and 61.74% for seven class on UTKFace and FairFace datasets. This simplification of the models decreases computational requirements in comparison to full-faced models. According to [46], the areas of facial skin are segmented, and dominant colors were added as features for classification. In Khan et al. (2021) they suggest that it is possible to use DCNNs for face segmentation with accuracy of 100%, 99.2%, 92%, and 93.2% over FERET and CASPEAL datasets, respectively. [40] used four handcrafted features (MLBP, HOG, color histogram, SURF-based) with an SVM classifier to classify ethnicity with an 89.14 per cent accuracy on UTKFace. Hamdi and Moussaoui [54] achieved a CNN ethnicity recognition performance of 80.22% on UTKFace by pre-training MobileNet and in combination with dropout layers to reduce overfitting and [28] used a pre trained MobileNet, reaching 80.22% on UTKFace. Kolla and Savadamuthu [39] investigate how the training data affects face recognition algorithms, but the authors demonstrate that if racial representation is equal, it won't erase bias; some facial qualities and racial features can compound it. The RFW dataset, ResNet 50 were used and 80% accuracy with respect to the ethnicities was achieved. In [47], both the linear and polynomial stacked attention mechanism models were introduced, and the Cupid dataset was used to test the Asian Face which reached 72.52% accuracy. According to [53], the Asian face dataset reaches 80.5% accuracy using transfer learning with VGG16. With the CJK dataset, [51] used E-CAE to accomplish this, and they were able to discriminate Asian ethnicities with an accuracy of 80.69%. The absence of Arab datasets was addressed by Umar Al-Humaidana et al. (2021), who created a labelled set of Arab datasets and used CNNs and deep clustering to achieve 56.97% accuracy.

Other models on geographically close race datasets, Samir Brahim et al., (2020) proposed hybrid approach of CNNs and Gabor filters with 92.5% accuracy. An example exists in [55] where they presented and proposed a deep convolutional neural network for Human Race Recognition (HRR) with an accuracy of 96.9%, outperforming the VGG and Inception V3 models. HHODTLF-FER, which is the fusion between VGG16, Inception v3, CapsNet, and BiLSTM with an accuracy of 99% in both VMER and UTKFace dataset, is proposed by [49] as a new best ethnicity recognition model. The CNN, GAN, ViT, AE models, and Ensemble CNN are developed by [35] for the purpose of ethnicity classification with an accuracy rate of 78.9% with Ensemble CNN. The authors also tested six pretrained models, however in some cases there were unbalanced datasets with one giant one white category of 42.51% and a difference in all the other classes. Among the suggestions of Greco et al. (2020), it is suggested that VGGFace2 Mivia Ethnicity Recognition (VMER) dataset consists of 3 million face images. For this dataset, the accuracy rate was 94.1% when the VGG-Face model was used. In [52], they created a large and balanced dataset of over 175,000 face images and trained four CNNs, of which Inception ResNet-v2 achieves 96.36% accuracy.

4.3.2 Interpretation and insights

Ethnicity recognition remains the most challenging among the three soft biometric tasks. Although CNN-based models achieve decent accuracy for majority ethnic categories (e.g., White, Asian), the performance degrades substantially for minority or mixed groups. This discrepancy stems from class imbalance in training datasets and inherent bias in model training. Furthermore, ethnicity labels are often self-reported or inferred subjectively, introducing annotation noise. These findings reveal that while current methods show promise, they risk reinforcing societal biases if not carefully validated. Future research must prioritize dataset balance and algorithmic fairness to ensure equitable recognition across all ethnicities.

4.3.3 Section summary

Across gender, age, and ethnicity recognition, the comparative results consistently demonstrate the superiority of deep learning models over traditional methods. However, fairness and demographic bias remain critical issues, especially in ethnicity and age classification. Furthermore, the lack of standardized evaluation protocols hinders cross-comparison. These insights highlight the need for ethically sourced, balanced datasets and unified benchmarking practices in future soft biometric research. These insights reveal that while soft biometric recognition systems have advanced significantly in gender classification, challenges remain more pronounced in ethnicity and age domains. This reinforces the importance of addressing demographic bias, a theme that continues in the discussion of open challenges in Sect. 5.

5 Challenges & future directions

The field of age, gender, and ethnicity classification in computer vision—particularly when leveraging transfer learning—faces several key challenges. Below, we outline these challenges, their implications, and potential strategies for resolution.

Model Bias: Model bias frequently originates from imbalanced or non-representative training data. This can perpetuate existing societal biases and result in the misclassification of minority groups. To address this, training datasets must be diversified and balanced across gender, age, and ethnicity. Furthermore, incorporating fairness-aware algorithms—such as adversarial debiasing—and conducting systematic bias audits using techniques like counterfactual testing can improve fairness and trustworthiness (Buolamwini & Gebru 2018).

- Model Generalization: Variations in cultural and demographic contexts limit the generalizability of trained models. Solutions include domain adaptation techniques and fine-tuning with local data. Federated learning can also enable inclusive model training without compromising user privacy, ensuring more robust performance across global populations.
- Transparency in Gender Classification: Increasing complexity in neural networks often diminishes model interpretability. To ensure accountability, explainable AI (XAI) techniques such as SHAP and Layer-wise Relevance Propagation (LRP) should be applied.

These approaches help visualize how specific features influence predictions, promoting trust in gender classification models.

- Ethical Concerns: The use of age, gender, and ethnicity as biometric traits raises ethical issues, particularly related to privacy, consent, and potential misuse. Ethical deployment mandates strong data governance frameworks, anonymization practices, adherence to data protection regulations (e.g., GDPR), and involvement of institutional review boards or ethics committees.
- Limited Diverse Datasets: A major limitation in building equitable models is the lack of diverse, high-quality labeled data. To resolve this, collaboration with international organizations and underrepresented communities is necessary. Synthetic data generation using Generative Adversarial Networks (GANs) may also assist in augmenting real-world diversity.
- Inconsistent Ethnicity Categories: Differing definitions and categorizations of ethnicity across studies hinder reproducibility and cross-comparison. Adopting standardized categories as proposed by international bodies such as the WHO or UN can foster greater consistency and comparative clarity.

Future research must extend toward deploying biometric models in emerging platforms such as Wireless Sensor Networks (WSNs) and IoT-enabled smart healthcare systems. These applications demand secure and real-time recognition frameworks. As emphasized in recent literature, soft biometric systems integrated with cryptographic protections and decentralized architectures could offer adaptive, resilient solutions for healthcare 4.0 and other cyber-physical applications.

6 Conclusion

This review paper discussed at length the potential applications of CNN models for age, gender, and ethnicity classification, recognizing its usefulness to fine-tune pre-trained models on different tasks. Transfer learning has demonstrated remarkable potential in facilitating the efficient transfer of knowledge from large to small datasets, as well as improving accuracy using feature extraction techniques. The papers discussed in this review largely concentrated on improving demographic identification with the use of random images, with many depending on databases of celebrity face images. Different pre-trained CNN architectures were investigated, with special interest in strategies such as hyperparameter fine-tuning to prevent overfitting, enhance generalization, and maximize accuracy. Emphasis was on the overarching importance of data quality and preprocessing cleaning, standardization, and enriching data being requisite to train models on informative features, fostering reliability in real-world applications. The literature surveyed further pointed out the importance of handling learning rates carefully while fine-tuning to achieve a balance which steers clear of overfitting as well as underfitting. This review collectively points to the tremendous potential that transfer learning must enhance demographic recognition tasks, but also to advance their deployment in diverse real-world applications.

Acknowledgements Not applicable.

Author contributions BH and PP gave the initial concept and idea of the review, AM provided suggestions to collect data and summarized it. While HS and SB investigated all the relevant sources and SRD and UK prepared the initial draft using the technical content produced by HS and SB. BH, PP and AY did revisions of the draft and produced the final version of manuscript.

Funding Not applicable.

Data availability Not applicable.

Declarations

Ethical approval This research does not involve any studies with human participants or animals performed by any of the authors. Therefore, ethical approval was not required.

Consent to participate Not applicable.

Consent to publish Not applicable.

Competing interests The authors declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

7 References

1. Nguyen T, Hoang D, Pham Q (2021) Gender classification with convolutional neural networks: improving performance with transfer learning and data augmentation. Int J Comput Vision 129(2):318–330
2. Hassan B, Fiaz M, Sherazi HHR, Butt UJ (2023) Annotated pedestrians: a dataset for soft biometrics estimation for varying distances. IEEE J Sel Top Signal Process 17(3):699–707. https://doi.org/10.1109/JSTSP.2023.3234494
3. Guo J, Mu H (2019) Age classification and estimation based on convolutional neural networks. Comput Vis Image Underst 179:28–38
4. Panis E, Galan G, Rousset M, Sebe N (2016) Age classification using convolutional neural networks: a comprehensive analysis. IEEE Trans Affect Comput 7(1):63–75
5. Geng X, Roij A, Hossain MA, Lim M (2018) A comprehensive study on age estimation using deep convolutional neural networks. Int J Comput Vision 126(1):44–61
6. Chen Z, Huang H, Zhang H, He X (2020) Age classification based on convolutional neural networks with data augmentation. J Vis Commun Image Represent 67:102753
7. Jain AK, Ross A, Nandakumar K (2020) Introduction to Biometrics. Springer Science & Business Media
8. Zhang N, Luo P, Loy CC, Tang X (2017a) Learning social relation traits from face images. Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp 3631–3639
9. Liu Z, Wen Y, Yu Z, Li M (2015) SphereFace: Deep hypersphere embedding for face recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 212–220
10. Zhang X, Yang Y, Luo P (2017b) UTKFace: A large-scale dataset for face recognition with age, gender, and ethnicity labels. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 4799–4807

11. Wang L, Zhang X, Zhang L (2019) Gender classification using deep convolutional neural networks. J Vis Commun Image Represent 63:102598
12. Wang M, Li X, Zhang Z (2019) FairFace: a benchmark dataset for ethnicity classification. IEEE Trans Image Process 28(8):3995–4008
13. Karkkainen H, Joo J (2021) Analyzing ethnic bias in facial recognition systems and datasets. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp 3227–3235
14. Liu Y, Tang X, Liu X (2022) Exploring dataset biases in ethnicity classification for facial recognition. IEEE Trans Pattern Anal Mach Intell 44(5):901–912
15. Albiero A, Cohn J, Zhai Y (2020) Fairness-aware algorithms for ethnicity classification: reducing bias in machine learning models. J Artif Intell Res 68(3):301–315
16. Hassan B, Izquierdo E (2022a) OneDetect: A federated learning architecture for global soft biometrics prediction. 2022 Int Conf Intell Syst Comput Vis (ISCV) [Preprint]https://doi.org/10.1109/iscv54655.2022.9806101
17. Hassan B, Izquierdo E (2022b) Rsfs: A soft biometrics-based relative support features set for person verification. In: Fourteenth International Conference on Digital Image Processing (ICDIP 2022) (Vol. 12342, p. 1234202). SPIE
18. Hassan B, Izquierdo E, Piatrik T (2021) Soft biometrics: a survey. Multimedia Tools Appl 83(5):15151–15194. https://doi.org/10.1007/s11042-021-10622-8
19. Hassan B, Izquierdo E, Piatrik T (2021b) Soft biometrics: a survey: benchmark analysis, open challenges and recommendations. Multimedia Tools Appl, 1–44
20. George G, Uppin C, Bello UA (2024) Human age Estimation from face images with deep convolutional neural networks using transfer learning [Preprint]. https://doi.org/10.20944/preprints202401.0350.v1
21. Chae Y et al (2023) Age prediction from face images via Contrastive Learning. 2023 18th International Conference on Machine Vision and Applications (MVA) [Preprint]. https://doi.org/10.23919/mva57639.2023.10216074
22. Kancharlapalli TP, Dwivedi P (2023) A novel approach for age and gender detection using deep convolution neural network. In: 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom). IEEE, pp 873–878
23. AlQadi RA, Batouche M (2023) Application of advanced deep learning techniques for face detection and age Estimation. 2023 1st Int Conf Adv Innovations Smart Cities (ICAISC) [Preprint]. https://doi.org/10.1109/icaisc56366.2023.10085089
24. Kanwar A, Singh KD (2023) Prediction of age, gender, and ethnicity using CNN and facial images in real-time. 2023 IEEE World Conference on Applied Intelligence and Computing (AIC) [Preprint]. https://doi.org/10.1109/aic57670.2023.10263824
25. Christopher MV et al (2023) Comparing age Estimation with CNN and EFFICIENTNETV2B1. Procedia Comput Sci 227:415–421. https://doi.org/10.1016/j.procs.2023.10.541
26. Haseena S et al (2022) Prediction of the age and gender based on human face images based on deep learning algorithm. Comput Math Methods Med 2022:1–16. https://doi.org/10.1155/2022/1413597
27. Gowroju S et al (2022) Deep Neural network for accurate age group prediction through pupil using the optimized unet model. Mathematical problems in engineering, 2022, pp. 1–24. https://doi.org/10.1155/2022/7813701
28. Kothari S, Deshmukh S, Mehta S (2022) Comparison of age, gender and ethnicity prediction using traditional CNN and transfer learning. 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT) [Preprint]. https://doi.org/10.1109/icccnt54827.2022.9984552
29. Hiremath JS et al (2022) Age detection based on facial features using support vector machine. 2022 International Conference on Knowledge Engineering and Communication Systems (ICKES) [Preprint]. https://doi.org/10.1109/ickecs56523.2022.10060119
30. Dagher I, Barbara D (2021) Facial age estimation using pre-trained CNN and transfer learning. Multimedia Tools Appl 80(13):20369–20380. https://doi.org/10.1007/s11042-021-10739-w
31. Garain A et al (2021) GRA_Net: a deep learning model for classification of age and gender from facial images. IEEE Access 9:85672–85689. https://doi.org/10.1109/access.2021.3085971
32. Agbo-Ajala O, Viriri S (2020) Deep learning approach for facial age classification: a survey of the state-of-the-art. Artif Intell Rev 54(1):179–213. https://doi.org/10.1007/s10462-020-09855-0
33. Gupta J, Pathak S, Kumar G (2022) Deep learning (CNN) and transfer learning: A review. J Phys: Conf Ser 2273(1):012029. https://doi.org/10.1088/1742-6596/2273/1/012029
34. Combs K, Lu H, Bihl TJ (2023) Transfer learning and analogical inference: a critical comparison of algorithms, methods, and applications. Algorithms 16(3):146. https://doi.org/10.3390/a16030146
35. Jayasekara BG, Hevapathige A (2022) Deep neural architectures for ethnicity classification in face images. 2022 Int Res Conf Smart Comput Syst Eng (SCSE). https://doi.org/10.1109/scse56529.2022.9905163. Preprint

36. Belcar D, Grd P, Tomičić I (2022) Automatic ethnicity classification from middle part of the face using convolutional neural networks. Informatics 9(1):18. https://doi.org/10.3390/informatics9010018
37. Abdulrahman BA, Mustafa NE (2024) Iraqi Kurd or Arab male authenticity detection based on facial feature. UHD J Sci Technol 8(1):64–77. https://doi.org/10.21928/uhdjst.v8n1y2024.pp64-77
38. Kalkatawi A, Saeed U (2024) Ethnicity classification based on facial images using deep learning approach. Int J Adv Comput Sci Appl 15(2). https://doi.org/10.14569/ijacsa.2024.0150223
39. Kolla M, Savadamuthu A (2023) The impact of Racial distribution in training data on face recognition bias: A closer look. 2023 IEEE/CVF Winter Conf Appl Comput Vis Workshops (WACVW) [Preprint]https://doi.org/10.1109/wacvw58289.2023.00035
40. Wirayuda TA, Munir R, Kistijantoro AI (2023) Compact-fusion feature framework for ethnicity classification. Informatics 10(2):51. https://doi.org/10.3390/informatics10020051
41. Abdulwahid AA (2023) Classification of ethnicity using efficient CNN models on morph and feret datasets based on face biometrics. Appl Sci 13(12):7288. https://doi.org/10.3390/app13127288
42. Makolo AU, Dada OV (2023) Ethnicity classification: a machine learning approach. Egypt Comput Sci J 47
43. Yang C-T et al (2023) Implementation of an attention mechanism model for facial beauty assessment using transfer learning. Diagnostics 13(7):1291. https://doi.org/10.3390/diagnostics13071291
44. Dhanaraj RS, Srid M (2023) Face warping detection and localization in a digital video using transfer learning approach [Preprint]. https://doi.org/10.21203/rs.3.rs-2950306/v1
45. Jain N, Kumar R, Goyal D (2023) A comparative study on face recognition using Deep Learning. Recent advances in sciences, engineering, information technology & management [Preprint]. https://doi.org/10.1063/5.0154462
46. Ridzuan SS, Omar Z, Sheikh UU (2023) Classifying ethnicity of the pedestrian using skin colour palette. Lecture Notes Electr Eng 105–116. https://doi.org/10.1007/978-981-19-8703-8_9
47. Li K, Zhang J, Shan S (2023) Self-supervised learning for fine-grained ethnicity classification under limited labeled data. 2023 IEEE 17th Int Conf Automatic Face Gesture Recognit (FG) [Preprint]https://doi.org/10.1109/fg57933.2023.10042748
48. Gan J et al (2022) Facial beauty prediction fusing transfer learning and broad learning system [Preprint]. https://doi.org/10.21203/rs.3.rs-1349480/v1
49. Obayya M et al (2022) Optimal deep transfer learning based ethnicity recognition on face images. Image Vis Comput 128:104584. https://doi.org/10.1016/j.imavis.2022.104584
50. AlBdairi AJ et al (2022) Face recognition based on Deep Learning and FPGA for ethnicity identification. Appl Sci 12(5):2605. https://doi.org/10.3390/app12052605
51. Ng WW, Zhou Z, Wang T (2021) Fine-grained facial ethnicity recognition based on dual convolutional autoencoders. 2021 11th International Conference on Intelligent Control and Information Processing (ICICIP) [Preprint]. https://doi.org/10.1109/icicip53388.2021.9642208
52. Darabant AS, Borza D, Danescu R (2021) Recognizing human races through machine learning—a multi-network, multi-features study. Mathematics 9(2):195. https://doi.org/10.3390/math9020195
53. Gao S et al (2020) Facial ethnicity recognition based on transfer learning from deep convolutional networks. 2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE) [Preprint]https://doi.org/10.1109/aemcse50948.2020.00073
54. Hamdi S, Moussaoui A (2020) Comparative study between machine and deep learning methods for age, gender and ethnicity identification. 2020 4th Int Symp Inf its Appl (ISIA) Preprint. https://doi.org/10.1109/isia51297.2020.9416549
55. AlBdairi AJ, Xiao Z, Alghaili M (2020) Identifying ethnics of people through face recognition: a deep CNN approach. Sci Program 2020:1–7. https://doi.org/10.1155/2020/6385281
56. Greco A (2020) Benchmarking deep network architectures for ethnicity recognition using a new large face dataset. Mach Vis Appl. https://doi.org/10.1007/s00138-020-01123-z

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Bilal Hassan[2] · Hasti Soudbakhsh[2] · Samaneh Bahrami[2] · Sonjoy Ranjan Das[3] · Preeti Patel[1] · Amanullah Yasin[4] · Usman Khan[4]

✉ Bilal Hassan
b.hassan@northumbria.ac.uk

Hasti Soudbakhsh
hasti.soudbakhsh@northumbria.ac.uk

Samaneh Bahrami
samaneh.bahrami@northumbria.ac.uk

Sonjoy Ranjan Das
sonjoy.d@esl.ac.uk

Preeti Patel
p.patel@londonmet.ac.uk

Amanullah Yasin
amanyasin.h11@bahria.edu.pk

Usman Khan
usmankhanakbar@gmail.com

1 School of Computing & Digital Media, London Metropolitan University, London, UK

2 Faculty of Engineering & Environment, Northumbria University London Campus, London, UK

3 Department of Computer Science, Elizabeth School of London, London, UK

4 Department of Computer Science, Bahria University Islamabad, Islamabad, Pakistan