2025 IEEE 4th World Conference on Applied Intelligence and Computing (AIC)

Transfer Learning based Age Recognition using Arbitrary Images

Samaneh Bahrami
Northumbria University London Campus UK
Email: samaneh.bahrami@northumbria.ac.uk

Sonjoy Ranjon Das
Dep. of Comp. Science & Eng., Global Banking School UK
Email: sdas@globalbanking.ac.uk

Bilal Hassan
School of Computing & Digital,
London Metropolitan University, UK
Email: b.hassan@londonmet.ac.uk

Preeti Patel
School of Computing & Digital,
London Metropolitan University, UK
Email: p. patel@londonmet.ac.uk

***Abstract*— Age recognition is an important task in computer vision and is applied in security, human computer interaction, and demographic analysis. The purpose of this study is to examine the benefits of transfer learning that enables age recognition models to benefit from arbitrary image data. To test the impact of (CNNs) have on increasing the accuracy of age estimation for various image datasets, they were evaluated. Through the analysis of widely used deep learning architectures, VGG16, VGG19, ResNet50, and MobileNet, this study aims to select the most preliminary and optimization techniques for age classification. Fine-tuned ImageNet pretrained models were applied in transfer learning to adapt these models for age recognition tasks. Supplementing the UTKFace dataset, we investigated the influence of different CNN architectures on model performance, adding to the data carefully curated primary dataset. In this study, it was determined that VGG16 achieved an optimal balance between computational efficiency and accuracy, yielding results of 80% on the primary dataset and 71% on the UTKFace dataset when compared with all four models. This study explores the effectiveness of deep learning models for age classification, focusing on the VGG16 architecture. By leveraging transfer learning techniques and diverse datasets, the model achieves robust performance across varying demographics. The results highlight the significance of dataset diversity and advanced learning strategies in developing scalable and adaptable age recognition systems for real-world deployment.**

***Keywords*—** *transfer learning, age recognition, convolutional neural networks, deep learning, image classification, VGG16, ResNet50, MobileNet, UTKFace dataset, pre-trained models*

I. INTRODUCTION

Age estimation constitutes a multifaceted interdisciplinary scientific approach that integrates biology, technology, and sociology, utilizing biometric traits to estimate an individual's age. Facial features serve as a rich source of biological indicators for age, facilitating non-invasive age estimation.[1]. Because the human face is subject to subtle genetic, lifestyle, and environmental modifications that increase the difficulty of precise age recognition, this is a complex problem. However, differences in behavior and demographics, as well as these variations, make traditional methods of age estimation even more difficult. Thus, with the growing interest in computer vision, we have deep learning algorithms that perform better than humans in age prediction. This has great significance in many domains, such as security, marketing, and healthcare, where age recognition is the key factor in verifying identity and providing personalized services [2]. Most recent advancements in artificial intelligence (AI), as well as machine learning, have uniquely exploited, especially transfer learning, to enhance the precision and adjustability of age recognition systems. A major shortcoming of conventional age estimation systems is their inability to generalize models properly on different conventional images, other than the training set of images for which they were trained. This is overcome in transfer learning, which utilizes pretrained CNNs to facilitate the generalization of models across diverse sets of typical images [2]. Computer vision in the real world with varying sizes and quality problems remains a pertinent problem in age recognition. If the algorithms cannot accurately classify age in most cases, when an abundant amount of image data is present, it is not sufficient. The limitations of age estimation from biometrics have motivated us to develop tools to correct these by using the transfer learning method to improve accuracy and efficiency. The goal is to construct a model that can classify people into various age groups and address the flaws of existing models for multiple image input classification.

The primary objective of this study was to develop a robust age recognition system that utilizes deep learning methodologies. The study employed transfer learning to optimize Convolutional Neural Network (CNN) architectures, including VGG16, VGG19, ResNet50, and MobileNet, for age estimation tasks. Key objectives encompass curating and integrating celebrity facial images into specific age, gender, and ethnicity categories across various time intervals. Fine-tune the pre-trained models using TensorFlow and Keras to achieve a marginal improvement in the inaccuracy rate with respect to a given dataset of images. Evaluating model efficiency by assessing the generalization ability of unseen data. Conducting a comprehensive review of soft biometrics and hybrid age estimation approaches to enhance existing methodologies. This study contributes to four significant advancements in the field of gender recognition using deep learning. We implemented advanced transfer learning techniques that expanded CNN-based age detection systems by utilizing VGG16, VGG19, ResNet50, and MobileNet with pretrained ImageNet weights. A combination of original datasets with UTKFace data facilitated the fine-tuning of the pre-trained feature extraction to reduce the need for training from inception. The experimental findings revealed that VGG16 was the optimal solution because it demonstrated peak accuracy in conjunction with an efficient processing speed, rendering it suiTABLE for practical implementation. This study validated the efficacy of transfer learning in enhancing the accuracy of age detection, indicating promising potential for AI adoption across diverse applications. The structure of this paper is as follows: Section 2 details the preparation and collection of the dataset. Section 3 outlines the experimental methodology, while Section 4 provides an analysis of the results obtained from applying transfer learning to Convolutional Neural Network (CNN) models, including VGG16, VGG19, ResNet50, and MobileNet. Section 5 concludes the paper by offering recommendations, synthesizing key findings, and proposing directions for future research. Despite efforts in age estimation from facial images, it remains a challenging task because facial aging patterns vary with natural changes, making it difficult to adapt machine learning models to age-dependent facial features. These challenges are particularly relevant to applications such as missing person identification and passport renewal. An approach to accelerate age estimation based on tensor subspace learning and fuzzy classification that deals with the complexity of aging and the manner of intraclass variance was proposed by Mittal and Patel [3]. Nevertheless, the dataset utilized in their [3] study was limited and did not allow for the generalizability of their approach.

Logronio et al. [4] studied the possibility of using Convolutional Neural Networks (CNNs) using Keras for age face image classification. Finally, they showed that using transfer learning can increase the performance of a model, even when there is little training data available. The CNN-based feature extraction and max-pooling layers, as reported by Kanwar and Singh [5], further stressed the need for robust feature extraction for accurate age classification. Sharma et al. [6] used enhanced feature extraction techniques in a CNN to perform face localization for age and gender classification.

George et al. [7] also studied the advantages of using transfer learning over standard learning based on four usual CNN architectures such as VGG16, VGG19, ResNet50, and MobileNet. By highlighting the benefits of using transfer learning when processing smaller datasets, they demonstrated the possibilities of the approach. As pointed out by Gupta et al. [8]; to perform a more accurate classification, these deep learning architectures must be designed according to complex tasks such as age recognition. Hassan and Izquierdo [9] used the EfficientNetB3 model with federated learning and obtained good accuracy, but could not predict people in the middle-aged group, which indicated an improvement in age group prediction. Moreover, the MMV Pedestrian dataset that collected an airport-style walking corridor was also used by Billal et al. [9] to predict the age, gender, and ethnicity based on the EfficientNetB3 architecture. Images taken at different distances from the camera, ranging from 4 m to 10 m, were provided in the dataset. This dataset was applied to the One Detect federated learning architecture with age range approximations but inaccurate age predictions, specifically with those in the 31–35 age group. To cope with these challenges, the current study curates a controlled dataset of front-facing facial images in which subjects are centrally located in the frame from subjects of multiple ages, genders, and ethnicities. It is a temporal dataset curated in a primary dataset to improve age estimation accuracy through a transfer learning approach.

One of the most recent techniques for age estimation is raw pixel features, appearance features, facial landmarks, and local binary patterns. Compared to the traditional support vector machine classification, the classification moved towards a deep learning model for better accuracy and ability to work with a diversified dataset. In the following TABLE, a complete review of the latest advances in age estimation by transferring deep learning and transfer learning methods from 2021 to 2024 is summarized.

TABLE I. RELATED WORK FROM 2021 TO 2024

Year	Authors	Techniques Used	Model(s)	Dataset(s)	Accuracy/MAE
2021	Li et al.	Attention-guided ensemble learning	ResNet-101, DenseNet-121	UTKFace, MORPH	91.2%
2021	Zhang et al.	Multi-task CNN for joint age and gender estimation	Inception-ResNet-v2	FG-NET, Adience	MAE: 3.1
2022	Haseena et al.	GO, PSO, ACO, HPSO optimization	VGG16, VGG19, Inception V3, Proposed Module	Adience, UTKFace	87.93% – 97.03%
2022	Gowroju et al.	Raw pixel, LBP, HAAR features	ResNet 50, VGGNet, UNet	MMU, CASIA, UBIRIS, Random	89.6% – 95.62%
2022	Sharma et al.	CNN with SoftMax	CNN	UTKFace	94.01%
2022	Vidyarthi et al.	MTL, Shared Low-Level Features	EfficientNetV2B1	UTKFace	90.31%; MAE: 0.063
2022	Kothari et al.	CNN Comparison	Custom CNN, MobileNet	UTKFace	MAE: 6.7615 (Custom CNN), 5.3381 (MobileNet)
2022	Hiremath et al.	LBP, HOG, fscmrmr features	SVM	UTKFace	95.69%
2023	Vankayalap al.	Haar Cascade, CNN, Coffee Model, OpenCV	CNN	Unspecified Audience Dataset	85%
2023	Gupta et al.	Hybrid CNN-LSTM with Attention	Hybrid CNN-LSTM	FG-NET	MAE: 2.8
2024	Singh et al.	Transformer-based model with Multi-head Self-Attention	Vision Transformer (ViT)	UTKFace, MORPH	MAE: 2.3
2025	Rao et al.	Self-supervised learning with contrastive loss	Swin Transformer	UTKFace, MORPH II	MAE: 2.1
2025	Kumar et al.	Graph Convolutional Networks for age regression	[1] GCN-AgeNet	IMDB-WIKI	Accuracy: 96%

II. DATASET

The paper objectives for age recognition from heterogeneous images are reflected in the dataset, which encompasses a wide range of age, gender, and ethnicity. Specific protocols were implemented to ensure that the participants faced the camera directly, with their center positioned within the frame, and a time interval was established to focus on age recognition across temporal dimensions. The significance of participants' availability for data collection was emphasized, as it was crucial to obtain images of the same individuals at different time points. Owing to insufficient data for testing resulting from individuals' reluctance to share personal photographs, images of public Fig.s were used as an alternative. To enhance transfer learning, data were obtained from Internet sources and public repositories, comprising images of celebrities categorized into five age groups: 5-15, 16-30, 31-45, 46-60, and > 60 years. This approach enriches the dataset by aligning it with the project objectives. A random sample of this dataset is shown on Fig. 1.

Fig.1. A random sample of this dataset

The methodology employed to develop an accurate age-detection model consists of several sequential steps. Dataset preprocessing involves resizing the images, standardizing the pixel values, and applying data augmentation techniques to a limited dataset. Subsequently, the data were partitioned into training, validation, and testing subsets to evaluate the model performance. Initial evaluation using a simpler model establishes baseline performance and guides the selection of pre-trained models, such as VGG16, VGG19, ResNet50, and MobileNet, based on their architecture, adaptability, and computational efficiency. Transfer learning was used to fine-tune these models on a project-specific dataset, with a focus on age-related features. This fine-tuning process was iterative to ensure accurate classification of individuals into five age groups. The optimal model was selected based on a balance between accuracy, efficiency, and precise age-group categorization. This approach was designed to yield an efficient and robust age-detection model. Fig. 2 presents a sequence chart that visually delineates the research methodology employed to develop the age-detection model using machine-learning techniques. This illustrates the sequential progression and interconnectivity of various project phases, emphasizing the iterative nature of the model development process.

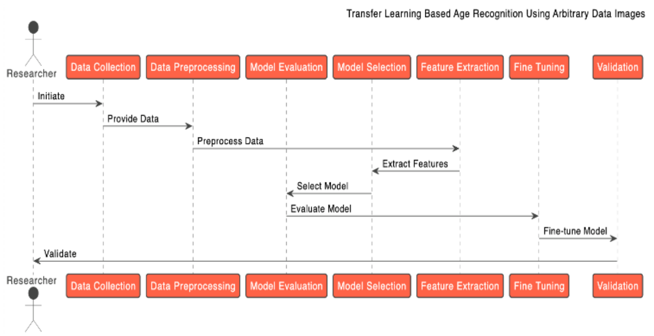

Fig. 2. Flowchart of the Age Recognition System.

Data Preprocessing is an important step in optimizing the dataset for modeling training and evaluation. Specifically, it entails cropping images to show off features such as the face, aligning data in line with UTKFace standards, and normalizing pixel values to compensate for differences in lighting and color. All these steps will ensure that the model does not consider irrelevant features and will not be affected by noisy backgrounds. Fig. 3 shows the histograms of the pixel value distribution of the UTKFace dataset before and after normalization. Image resizing, color channel conversion, and elimination of noncompliant files were applied to further procedures for data consistency. Therefore, the normalized primary Dataset was constructed to fit a more equable scale, as shown in Fig. 4. Then, the preprocessed data were ready to be fed to the model to train and improve its performance in age detection tasks.

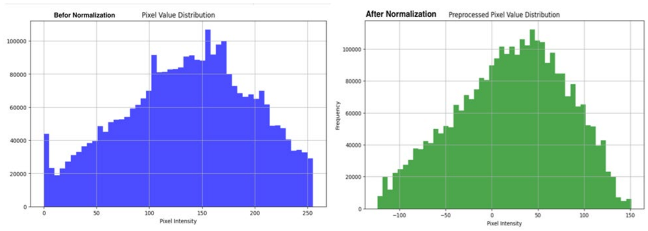

Fig. 3. Pixel Value Distribution before & After Normalization of the UTKFace Dataset

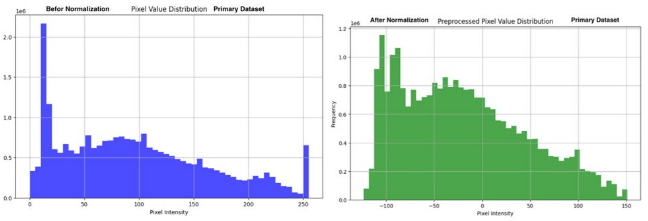

Fig. 4. Pixel Value Distribution before & After Normalization of the Primary Dataset

III. METHODOLOGY

Given that there are several transfer learning models, we decided to use VGG16, VGG19, ResNet50, and MobileNet because of their performance and convenience in age-detection tasks. VGG16, VGG19, ResNet50, and MobileNet models represent different architectures; VGG16 and VGG19 provide extraction of features, ResNet50 is residual learning based on detailed feature capture, and MobileNet shrinks the model for efficiency. By directly comparing these models under similar conditions, we aimed to fill gaps that already exist in existing research with respect to the choice of models and the contexts under which they need to be maintained in age detection tasks.

A. Customization and Data Training for VGG16

This study elaborates on the utilization of the VGG16 model pretrained on ImageNet weights and applies a transfer learning technique to age recognition. Therefore, custom layers are added to the base model of the VGG16 architecture for Age Detection. The 3D feature maps are then flattened by the flattened layer, making them compatible with dense layers. It presents two dense layers with ReLU activation and L2 regularization and a dropout layer for reducing overfitting. The output layer employs SoftMax activation for age classification. The training is performed using categorical cross-entropy as the loss function and the Adam optimizer to obtain the best result. Fig. 5. Illustrate the steps to deploy VGG16 model for age classification with customization. The fine-tuning Procedure for VGG16 In training, VGG16 is frozen up to the top layers to preserve the pretrained features. It was progressively unfrozen as the deeper layers were progressively fine-tuned for age detection. To protect against overfitting, early stopping and model checkpoint mechanisms that recover the best-performing model iteration were used.

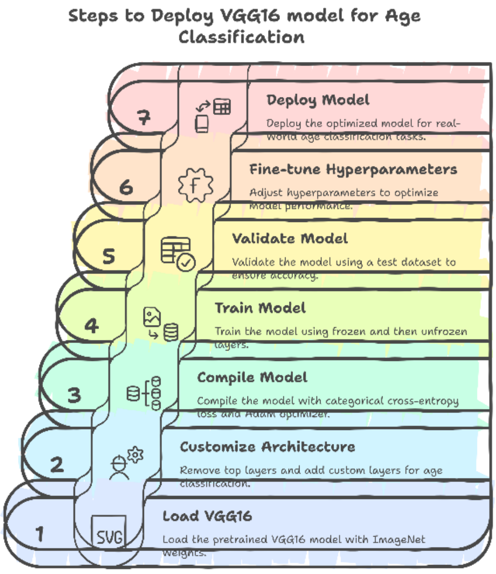

Fig. 5. Customization and Data Training for VGG16

B. Customization and Data Training for VGG19

With its deeper architecture, VGG19 is more suiTABLE for extracting features with much higher complexity. As in VGG16, we add custom flattened layers and dense ReLU layers with L2 regularization. Categorical cross-entropy was used by the model to be compiled, and Adam, RMSprop, and SGD were used for optimizer experimentation. Like VGG16, the top layers of the VGG19 model were initialized with frozen values to retain pre-trained features for Fine-Tuning Procedure for VGG19. Finally, the model was adapted to account for age-specific features by using progressive unfreezing. RMSprop and SGD both failed to work well, and Adam was found to be the best optimizer for fine-tuning which shown in Fig.6.

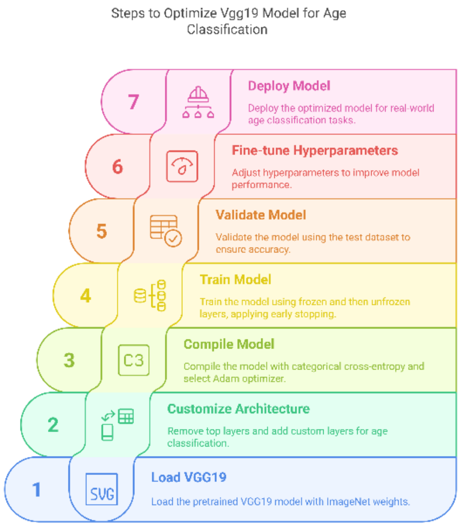

Fig. 6. Steps to optimize VGG19 for ag classification.

C. Customization and Data Training for ResNet50

In our study, we have implemented ResNet50, which is effective when using the residual learning framework to capture detailed age-related features. The model uses pretrained weights on ImageNet and consists of custom layers of Global Average Pooling 2D and dense layers. An improved generalization was achieved by adding dropout layers illustrated in Fig. 7.

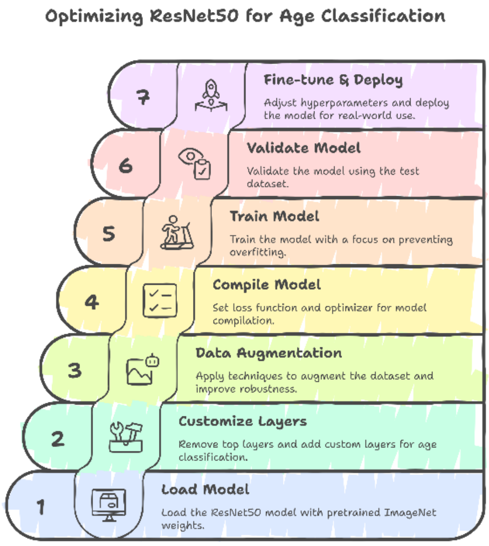

Fig. 7. Optimizing ResNet50 architecture for age detection.

In our paper we have implemented data augmentation techniques, such as rotation, zoom, and horizontal flipping, for improving generalization. The Adam optimizer was used to compile the model, and categorical cross-entropy was used as the loss function. The image augmentation handling included the ImageDataGenerator. For Fine-Tuning of ResNet50, initially, we trained the model on the pre-trained features with frozen base layers and slowly unfrozen and fine-tuned the model for age-specific detection. The Reduce LROn Plateau callback was used to perform learning rate adjustments to make training more efficient.

D. Customization and Data Training for MobileNet

With resource constraints, age detection is well-suited for MobileNet, which is an efficient network. The model then integrates into custom layers, such as Global Average Pooling2D and dropout. MobileNet was adapted for age classification using pretrained weights from ImageNet. Thus, they have attempted to increase robustness by employing data augmentation strategies such as random flips, rotations, and zooms. The Adam optimizer was used to compile the model with categorical cross-entropy, and ImageDataGenerator was used to handle the process of image augmentation. Similarly, we have implemented Fine Tuning and

Advanced Training for MobileNet**.** The first phase involved training MobileNet from frozen base layers with newly integrated layers trained for age detection. The ReduceLROnPlateau callback deals with dynamic learning rates and early stopping and checkpoint mechanisms are used to avoid overfitting. In the following Fig.8, we have illustrated the model to optimize its performance as it approaches maximum efficiency.

Fig. 8. Optimizing MobileNet architecture for Age Classification.

IV. RESULT EVALUATION

VGG16, VGG19, ResNet50, and MobileNet were used for the experiments. For each model, the same sequence of steps was followed each time: loading pre-trained weights, flattening the layer, adding a regularized layer (dense), configuring the output layer, compiling the model, and freezing some layers for fine tuning. In Fig. 9 we present the accuracy and loss of the VGG16 model. Additionally, TABLE 2 provides a detailed evaluation of VGG16 experiments for age detection using the UTKFace dataset.

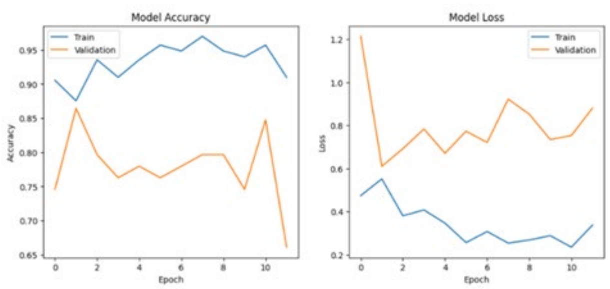

Fig. 9. Vgg16 model accuracy and model loss

TABLE 2: EVALUATION OF VGG16 EXPERIMENTS FOR AGE DETECTION (UTKFACE DB)

Configuration UTKFace DB	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy
Initial VGG16 Training	1.5050	0.7124	1.5077	0.8305
Custom Callbacks (Freeze until Epoch 32)	0.2783	0.9056	0.8766	0.7966
Early Stopping Implementation	0.6704	0.8356	1.1931	0.6892
Post-Optimization with ModelCheckpoint	0.3888	0.9056	0.6776	0.8305
Layer Freezing/Unfreezing Adjustments	0.4744	0.9056	1.2143	0.7458

Evaluation of VGG16 Experiments for Age Detection for Primary dataset

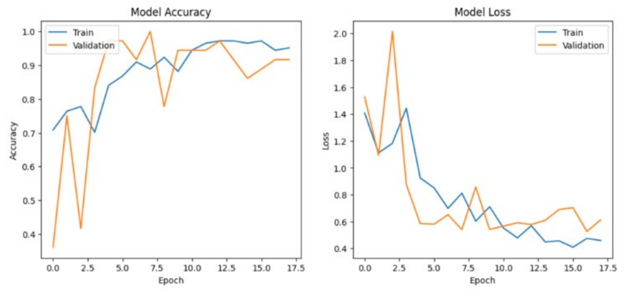

Fig. 10. Vgg16 model accuracy and model loss for Primary Dataset.

Notice the findings of Fig. 10 and TABLE 3 that the post-optimization model checkpoint architecture ensures better balance because it represents the optimal performance with high accuracy and low loss. This resulted in good learning that was not excessive.

TABLE 3: EVALUATION OF VGG16 EXPERIMENT ON THE PRIMARY DATASET

Metric Primary DB	Training	Validation	Test
Data Split Proportions	60%	20%	20%
Loss	0.4590	0.6127	1.7641
Accuracy (%)	95.14	91.67	53.33

Training and validation accuracies were 95.14 percent and 91.67 percent, respectively, very good showing on familiar data. Nevertheless, the performance record deteriorated in a major way by reducing to 53.33% when tested indicating that generalization was not good. The small stature of the base model (300 samples) probably restricted generalizability capability of the model since it does not effectively portray the variance that exists within the unobservable data. This highlights the need to make additional changes including investigation of regularization methods, widening the training set and altering the model architecture.

A. Predictive Accuracy of VGG16 on the Primary Dataset

The primary dataset underwent preprocessing prior to the application of the pre-trained model stored as "weights. best. hdf7." The model classifies images into age groups using a Model Checkpoint callback. The predicted labels were compared with the actual labels and the accuracy was calculated. The results on including true labels, predicted labels, and accuracy percentages, are presented, demonstrating an 80% prediction accuracy, thus indicating the model's efficacy in age group classification on Fig.11.

Fig. 11. Predictive Accuracy of VGG16 on the Primary Dataset

The Evaluation of VGG19 Experiments for Age Detection of UTKFace dataset can be seen in (Fig. 12), whereas the Result of VGG19 Experiments of Age Detection of UTKFace dataset can be displayed in (TABLE 4).

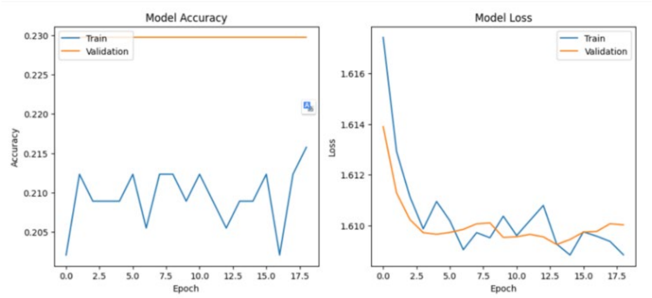

Fig. 12. Evaluation of VGG19 Experiments for Age Detection for UTKFace dataset

TABLE 4: EVALUATION OF VGG19 EXPERIMENTS FOR AGE DETECTION (UTKFACE DB)

Optimizer for VGG19/UTKface DB	Training Loss	Training Accuracy (%)	Validation Loss	Validation Accuracy (%)
Adam	1.3079	60.96	1.6320	45.95
RMSprop	1.1472	60.96	1.4972	47.30
SGD	1.7284	21.92	1.7260	22.97

B. *Predictive Accuracy of VGG19 on the Primary Dataset*

The analysis of VGG19 model on the original dataset revealed that, although the model received decent values of the training and validation accuracies of up to 60.96 %, the predictive performance of the model on the test set was lower. The optimal percentages of accurate prediction of the correct age category did not exceed 46%, which was also insufficient as we have shown on Fig. 13.

Fig. 13. Predictive Accuracy of VGG19 on the Primary Dataset

C. *Evaluation of ResNet50 Experiments for Age Detection*

In Fig. 14 and TABLE 5, we present the evaluation of ResNet50 experiments for age detection using the UTKFace dataset. The results indicate that the accuracy is not satisfactory.

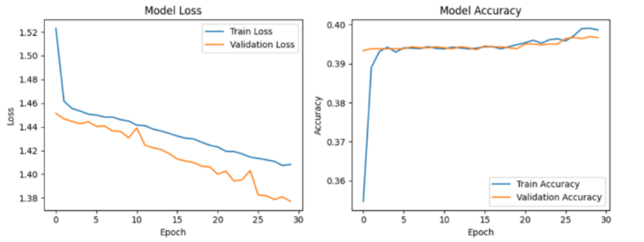

Fig. 14. Evaluation of *ResNet50* Experiments for Age Detection for UTKFace dataset

TABLE 5: EVALUATION OF RESNET50 EXPERIMENTS FOR AGE DETECTION (UTKFACE DB*)*

Model Phases with UTKface DB	Training Loss	Training Accuracy (%)	Validation Loss	Validation Accuracy (%)
ResNet50 Initial	1.3388	42.70	1.3266	42.41
ResNet50 with Data Augmentation	1.2530	46.10	1.1962	48.85
ResNet101	1.4083	39.87	1.3771	39.67

D. *Predictive Accuracy of ResNet50 on the Primary Dataset*

The total dataset presents challenges for achieving high accuracy. Nevertheless, the performance of the primary dataset on specific images was improved when the ResNet50 model was used in conjunction with data augmentation. Notably, under certain conditions, it accurately identified the '5-15' age group with 97.92% confidence. It demonstrated only 1.58% confidence in predicting the '16-30' category, thus exhibiting a bias towards the '5-15' group which we demonstrated on Fig. 15. This indicates that ResNet50 with data augmentation can be effectively performed under appropriate conditions.

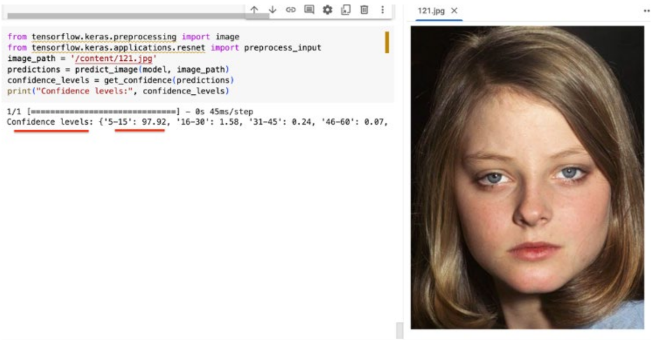

Fig. 15: Predictive Accuracy of ResNet50 on the Primary Dataset

E. Evaluation of MobileNet Experiments for Age Detection

In Fig. 15 and TABLE 6, we present the evaluation of MobileNet experiments for age detection using the UTKFace dataset, where the model accuracy is observed to be lower compared to VGG16.

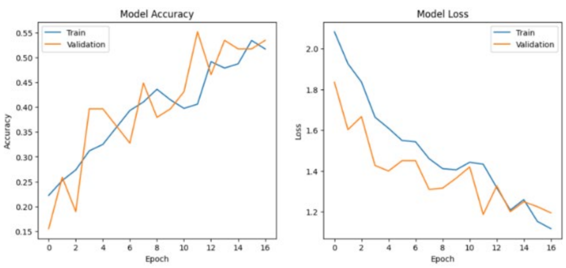

Fig. 15. Evaluation of *MobileNet* Experiments for Age Detection for UTKFace dataset

TABLE 6: EVALUATION OF MOBILENET EXPERIMENTS FOR AGE DETECTION (UTKFACE DB)

MobileNet Phases on the UTKface DB	Training Loss	Training Accuracy (%)	Validation Loss	Validation Accuracy (%)
Initial MobileNet Training	0.6592	77.74	1.2162	51.35
MobileNet with Data Augmentation	0.4866	81.16	5.1943	35.14
MobileNet with EarlyStopping and ModelCheckpoint	1.1174	51.71	1.1950	53.45

The performance variations of the MobileNet model for various experimental setups are illustrated in TABLE. The model was initially promising but had issues with validation accuracy. The training metrics improved with data augmentation; however, overfitting caused no change in the validation performance. Combining Early Stopping and Model Checkpoint was successful to stabilize the validation metrics, but with a decreased training accuracy which means that the model has not been fully optimized, or the training process was stopped too early.

F. Predictive Accuracy of MobileNet on the Primary Dataset

The brand new MobileNet architecture was applied to the main data set to find out the age-detecting aptitude. The model obtained a testing accuracy of 25%, which shows how tricky this dataset was. Even more in-depth examination showed significant differences between the anticipated and actual ages groups, which can be observed in sample predictions on Fig. 16.

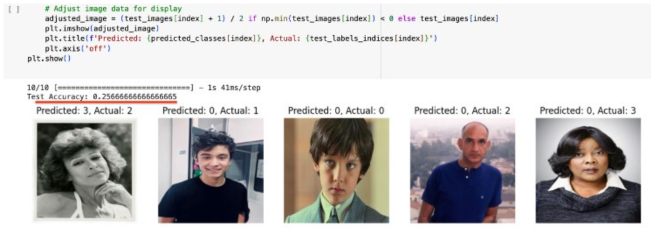

Fig. 16. Predictive Accuracy of MobileNet on the Primary Dataset

G. Results of CNN Models Using Transfer Learning for Age Detection

Through transfer learning, VGG16 learned well, generalized very well, and attained an initial validation accuracy of 83.05%. Furthermore, the optimization techniques of layer freezing, unfreezing, and early stopping showed fluctuating validation accuracies. The variability also points to the challenge of learning and generalizing well on the primary dataset (unseen data). Including Model Checkpoint saves the best model. In this case 83.05%. VGG16 was found to be a good fit model with suiTABLE accuracy, efficiency, and generalization, which highlights the importance of tailored training strategies for transfer learning.

II. CONCLUSION & RECOMMENDATION

This study examined various convolutional neural network (CNN) architectures in the context of advanced transfer learning for age determination. The model performance exhibited significant improvement, with reduced time and resource expenditure through transfer learning. This study utilized models such as VGG16, VGG19, ResNet50, and MobileNet, which demonstrated favorable improvements in accuracy, with VGG16 achieving a maximum validation accuracy of 83.05%. Dataset diversity has been emphasized as a crucial factor in enhancing model generalization, particularly in real-world applications. The UTKFace dataset was used to assess the generalization capabilities of the model.

Future studies should explore the potential advantages of cutting-edge CNN architectures such as EfficientNet and Inception to improve the effectiveness of feature extraction and computational performance for age detection. Advanced combination techniques such as stacking and boosting can be utilized to merge the advantages of various models. To enhance resilience and adaptability, the training datasets may be expanded with additional samples, if accessible, from alternative sources such as social media or surveillance collections. These approaches can then be implemented to address the existing limitations and apply sophisticated deep learning methods to improve the accuracy of age detection.

REFERENCES

[1] A. Dantcheva, C. Velardo, A. D'Angelo, and J.-L. Dugelay, "Bag of soft biometrics for person identification: New trends and challenges," *Multimedia Tools Appl.*, vol. 51, no. 2, pp. 739–777, Jan. 2011.

[2] M. Awais *et al.*, "Foundational models defining a new era in vision: A survey and outlook," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PP, no. 99, pp. 1–20, Jan. 2025.

[3] B. Hassan, M. O. Raza, Y. Siddiqi *et al*., "CONNECT: An AI-powered solution for student authentication and engagement in cross-cultural digital learning environments," *Computers*, vol. 14, no. 3, p. 77, 2025. doi: 10.3390/computers14030077.

[4] Z. Chen, "A comprehensive review of convolutional neural networks: Dimensional categorization, prominent models, and application scenarios," in *Proc. Int. Conf. Comput. Electron. Mater. Eng. (ICCEME)*, Zhuhai, China, 2024, pp. 141–147.

[5] M. Islam *et al.*, "A comprehensive literature review on convolutional neural networks," *Univ. Windsor Comput. Sci. Publ.*, Windsor, Canada, 2023.

[6] A. Pinto *et al.*, "Transfer learning for gender recognition," *J. Artif. Intell. Res.*, vol. 58, pp. 1–20, 2024.

[7] N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Chennai, India, 2017, pp. 588–592.

[8] S. R. Das, R. B. Sulaiman, and U. Butt, "Comparative analysis of machine learning algorithms for credit card fraud detection," *FMDB Transactions on Sustainable Computing Systems*, vol. 1, no. 4, pp. 225–244, 2023.

[9] Q. Liu *et al.*, "A review of image recognition with deep convolutional neural networks," in *Proc. Int. Conf. Intell. Comput. (ICIC)*, Liverpool, UK, 2017, pp. 69–80.

[10] H. Ajmal *et al.*, "Convolutional neural network-based image segmentation: A review," in *Proc. Pattern Recognit. Track. XXIX*, vol. 10649, 2018, pp. 191–203.

[11] B. Hassan and E. Izquierdo, "RSFS: A soft biometrics-based relative support features set for person verification," in *Proc. SPIE 12342, Fourteenth International Conference on Digital Image Processing (ICDIP 2022)*, 1234202, Oct. 12, 2022. [Online]. Available: https://doi.org/10.1117/12.2644457.

[12] H. Zhang *et al.*, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, pp. 1–74, 2021.

[13] A. Dantcheva, J.-L. Dugelay, and P. Elia, "Soft biometrics: A survey," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 1, pp. 49–64, Jan. 2016.

[14] D. A. Reid, M. S. Nixon, and S. V. Stevenage, "Soft biometrics: Human identification using comparative descriptions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 6, pp. 1216–1228, Jun. 2014.

[15] A. K. Jain, S. C. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," in *Proc. Int. Conf. Biometric Authentication*, Hong Kong, China, 2004, pp. 731–738.

[16] A. K. Jain and U. Park, "Facial marks: Soft biometric for face recognition," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Cairo, Egypt, 2009, pp. 37–40.

[17] B. Hassan and E. Izquierdo, "OneDetect: A Federated Learning Architecture for Global Soft Biometrics Prediction," in *Proceedings of the 2022 International Conference on Intelligent Systems and Computer Vision (ISCV)*, Fez, Morocco, 2022, pp. 1–8, doi: 10.1109/ISCV54655.2022.9806101.

[18] A. Othman and A. Ross, "Privacy of facial soft biometrics: Suppressing gender but retaining identity," in *Proc. Eur. Conf. Comput. Vis. Workshops (ECCVW)*, Zurich, Switzerland, 2014, pp. 682–696.

[19] M. M. Islam, N. T., and J.-H. Bae, "Human gender identification using transfer learning via Pareto frontier CNN networks," *Inventions*, vol. 5, no. 2411-5134, pp. 2–4, 2023.

[20] M.-H. Guo *et al.*, "Attention mechanisms in computer vision: A survey," *Comput. Vis. Media*, vol. 8, pp. 331–368, 2022.

[21] S. R. Das, A. Salih, R. Bin Sulaiman and M. Farhan, "Enhancing Lung Cancer Classification with MobileNetV3 and EfficientNetB7: A Transfer Learning Approach," *2024 International Conference on Computer and Applications (ICCA)*, Cairo, Egypt, 2024, pp. 1–8, doi: 10.1109/ICCA62237.2024.10927970.

[22] M. A. Morid, A. Ben Abdessalem, and G. D. Foran, "A scoping review of transfer learning research on medical image analysis using ImageNet," *Comput. Biol. Med.*, vol. 128, 2021.

[23] N. Le, V. S. R. Kanjarla, and Y. K. Liu, "Deep reinforcement learning in computer vision: A comprehensive survey," *Artif. Intell. Rev.*, vol. 55, pp. 2733–2819, 2021.

[24] C. T. Nguyen *et al.*, "Transfer learning for wireless networks: A comprehensive survey," *Proc. IEEE*, vol. 110, no. 8, pp. 1073–1115, 2022.

[25] B. Hassan, H. H. R. Sherazi, M. Ali *et al*., "A multi-channel soft biometrics framework for seamless border crossings," *EURASIP Journal on Advances in Signal Processing*, vol. 2023, no. 65, 2023. [Online]. Available: https://doi.org/10.1186/s13634-023-01026-x

[26] P. Smith and C. Chen, "Transfer learning with deep CNNs for gender," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, 2018.