Transfer learning based Human Face Traits Recognition in Individual Biometric Validation

Kiruthiga K[1*], *and Uma* maheswari K[2]

[1]Assistant Professor, Department of Artificial Intelligence and Machine Learning, KPR Institute of Engineering and Technology, Coimbatore, India
[2]Professor and HoD, Department of Information Technology, PSG College of Technology, Coimbatore, India

Abstract. The key component of contemporary biometric security systems is the accurate identification and validation of individuals using facial recognition technology. Advanced security systems can now be personalized through improved integration of gender recognition in their development. The face recognition models VGG16, VGG19 and ResNet50 together with other identifiers face issues to adapt to variations in skin tone specifically affecting users with darker skin complexion. The ability of these models strongly depends on skin melanin levels, which results in higher cases of false identification or bias towards users with these characteristics. Progress has been made, but the deficiency of skin-type representation during training con- tinues to result in performance discrepancies among numerous operational systems. The common method for evaluating and training face recognition models is based on machine learning technology. These face recognition datasets must have large-scale deployments of wide skin tone diversity and multiple age ranges to improve model recognition capabilities independently of demographic features. Transfer learning solutions create better model accuracy along with fairness, since they help large dataset- trained models adapt to small heterogeneous datasets. This method proves beneficial in the identification of human facial features for biometric validation because it showcases enhanced performance and shorter training periods, as well as universal real-world scenario applicability. The proposed work is suitable for practical real-world applications in Airport security, Border Control along with Border Surveillance and Law enforcement and finance, Healthcare, Mobile devices and security systems. In order to ensure more reliable and inclusive face recognition solutions, biometric systems can use transfer learning to overcome the challenges posed by diverse facial characteristics and skin tones. The results obtained from the experimental evaluation carried out on the UTKFace dataset show that the best performance is achieved by the ResNet50V2 model, which has the highest accuracy of 84% in the validation set compared to the accuracy achieved by the other architectures.

* Corresponding author: kiruthiga@kpriet.ac.in

1 Introduction

Traditional biometric systems include the assessment of singular biometric traits, such as fingerprint and iris recognition; these assessments are often invasive and require users to take an active role. As 'soft' biometric identifiers, facial traits satisfy the demand for system performance and security, as well as non-invasiveness (i.e, user acceptance of using their face for identification). By integrating AI into biometrics, there is now much greater availability and success of using AI for facial recognition and stores of human faces, breaking down the barriers of traditional machine learning techniques to correctly classify individual characteristics such as age, gender, ethnicity, etc. One of the most important characteristics of the human face for use with facial recognition systems is that the human face has unique features and is non-invasive. Facial recognition systems have the capability to recognize an individual based on the inherent features of the human face despite the environmental conditions (such as occlusion, changing light levels) under which the individual is being captured for use with the facial recognition system. Traditional machine learning algorithms and shallow neural network architectures have limitations in recognizing the differences in the features of the faces of humans belonging to different ethnic backgrounds. However, deep learning has achieved great advancements in the field of feature extraction. There are a lot of advantages associated with the application of deep learning in the field of biometric systems. Apart from the computational cost associated with the training of the deep neural network model, the application of the transfer learning mechanism may be viewed as having a second advantage in the sense of reducing the computational costs based on the knowledge gained from the training of the deep model.

This paper discusses a new approach for implementing transfer learning strategies in recognizing facial traits from images in developing personal biometric verification systems. This approach employs pre-trained convolution neural networks in identifying significant facial traits with an emphasis on resource efficiency in computing required by the extraction of these important traits in images. By integrating these strategies in a real-time solution, it becomes possible to identify significant biometric traits of importance such as gender and age in individuals. Based on these important properties and benefits, it becomes applicable in a commercial set-up as well as in any other set-up requiring high levels of security. Applying transfer learning strategies and human facial traits recognition enables an effective solution that is easily scalable and more cost-effective than previous solutions allowing for high levels of precision and interpretability in biometric verification solutions.

The UTKFace dataset was identified to be an established benchmark in the area of Face Attribute Prediction. This model described in this research has been trained and tested using this particular UTKFace dataset. The UTKFace Dataset contains 20,000 + labelled headshots that have been annotated with AGE, GENDER and ETHNICITY. The dataset contains images of individuals from 0 - 116 years old, and captures images that are similar to those found in the real world (e.g., lighting conditions, angles, facial expressions). The images are stored as 200x200 pixel RGB images. With its significant number of annotations, UTKFace is an outstanding resource for evaluating how well Transfer Learning will be applied to biometric systems, in both Binary (i.e., Gender Classification) and Multi-Class (i.e. Age Classification) tasks. However, many current systems do not evaluate fairness by gender or ethnicity, resulting in inaccurate predictions and biased results.

Metrics, including accuracy, precision, recall, F1-score, and training efficiency, will be used to thoroughly evaluate how well a given model(s) performed throughout the course of conducting this research. In addition, this research explores the interpretability of the predictions generated by the developed model(s) through visualization or "feature activation maps." This study emphasizes the generalization ability of its models across a variety of demographic groups to assess how effectively they can be used to develop reliable systems that are not only capable of achieving high levels of accuracy but also making AI decision-making transparent as to help build trust within the field of biometric applications that relate to individuals in particular. Subsequent sections in this report describe the design and training processes used to create the developed models as well as the theoretical foundation for transfer learning as applied in the present experiment to provide evidence supporting the concept of a reliable system for identifying facial traits. The objective of this report is to create scalable, interpretable, and ethically responsible AI-derived validation systems for biometric applications that may be utilized for demographic analysis and identity (verification) comparisons in real life situations.

Unlike many previous studies focusing only on the analysis of a single task in facial analysis, this research examines the concept of a multi-task transfer learning approach in which the prediction of age and gender is carried out at the same time using the CNN model. In this research a comparative analysis of several state-of-the-art architectures is also carried out in the context of the same training protocol, focusing on the effectiveness of the ResNet50V2 model in biometric facial trait recognition.

2 Research Objective and Contributions

The current study outlines the robust and flexible deep learning framework, which has been created to have the ability to detect human facial traits for biometric validation thereof. The transfer learning and the multitask learning methods are used to attain the reliability of the proposed system to have the ability to perform equally well under a wide range of demographic settings, besides being computationally efficient.
The major contributions of the proposed study are as follows:

- The proposed study has created a deep learning framework through the transfer learning method for the recognition of facial traits, such as age estimation and gender detection.
- The proposed study has created a multi-output CNN model for the ability to simultaneously carry out age prediction and gender detection.
- The proposed study has conducted a comparative study of the performance of the pre-trained CNN architectures, such as VGG16, ResNet50V2, ResNet101, EfficientNetB0, and EfficientNetB7.
- The proposed study has evaluated the performance of the proposed model through the UTKFace dataset, based on the accuracy, precision, recall, MAE, MSE, etc.
- The proposed study has identified the most efficient CNN architecture through the transfer learning method for the biometric validation of facial traits.

3 Literature Review

The article by Dey et al. [1] focuses on strategies of deep learning used to estimate the age and gender of a person and is a significant contribution to the state of the art in facial feature detection. The effectiveness of improvement in performance achieved through transfer learning has been illustrated by the authors in their evaluation on various CNN models. Conclusion: It is able to tackle the issues in age and gender estimation on various datasets using deep-learned features extracted by VGG or ResNet models. This will lay a foundation for age and gender estimation in more complex systems.

Khairnar et al. [2], the team of scientists, carried out an extensive study on the AI approaches that recognize the authenticity of real face images of real people who are also real and living, along with the significant implications of reliable biometric authentication. This research work identifies available avenues to be explored and concludes that the use of a deep learning model containing both texture and temporal features minimizes the chances of spoofing. The outcome of this research work is very significant to improve the security model of facial feature recognition systems.

An example of biometric authentication technology that utilizes the services of a deep learning algorithm for identifying the identities of users based on facial landmarks is shown in the work by Jie et al. [3]. Though their technology relies on intricate patterns created by spatial relationships among different parts of the human face for facilitating the task of separating genuine images from forged ones, it could have been an appropriate application for real-life situations that involve the real-time verification of an identity. The different parts of their methodology show the effective combination of geometric analysis for the human face with the deep learning representation system.

The authors Gawate and Dhote [4] demonstrate a multi- face recognition system combined with transfer learning to detect both liveness and presence in real time. The use of pre-trained models increases accuracy for these systems while greatly reducing the time required to train new users. This practice is especially beneficial when developing a large number of systems due to the need for fast and accurate identity verification. Their research provides evidence that using transfer learning in multi-face applications meets the requirements of today's biometric identity validation systems.

Proposal of a new method for generating deceitful facial images aiming to detect the liveliness of tampered facial images that simulate an actual individual was done by Long et al. [5]. The authors offer a method wherein both the generation of image features and the acceptance of the simulated facial image take place simultaneously using adversarial training with the use of generative adversarial networks (GANs). The methodological approach developed can deal with the differences among various spoofing attacks on the identity of an individual including new ones that have not been observed before, hence improving the biometric system authenticity.

Raman et al. [6] have proposed a system which involves computer-based deep learning algorithms to classify facial images into various age groups as per gender. The results obtained in their paper tend to justify that in ”real world” scenarios, systems which are designed with a focus on various demographics perform better compared to systems which do not focus on demographic aspects in their design. Further, accuracy in ascertaining age of an individual is enhanced when age is determined with regard to particular aspects as per gender. Hence, there is relevance to model structures designed specifically for the analysis of human facial features by these researchers.

A thorough research work by ElKarazle et al. [7] offers information about the techniques of facial age estimation based on deep learning as well as traditional machine learning. The authors have discussed regression techniques, feature engineering, as well as full deep learning models, underlining the importance of deep learning as the most universal technique, which should therefore be used as the primary technique for the estimation of facial age. ElKarazle et al. assert that their research work is a significant reference to the recognition of age based on face recognition, in order to determine the structure of the model as well as the training processes.

An Extreme Learning Machine (ELM) combined with deep learning algorithms was used by Michael and Shankar et al. in an attempt to predict age and gender in [8]. Their combination of rapid ELM training times with the capability to extract deep features using deep learning makes this style of a hybrid model particularly advantageous. Their experimental results indicate that applying hybrid techniques in this manner can improve both speed and accuracy, enabling them to be suitable for applications that demand real-time biometric authentication.

Tilki et al. [9] conducted an evaluation of convolutional networks across a number of different datasets with a focus on gender classification using deep learning methods. In their work on gender classification using deep CNN's, the authors highlight that these networks are relatively resilient to common variables seen in the real world (e.g. changes in lighting, orientation, and facial expression) that are typically encountered by researchers when working with real-world problems. The authors conclude that deep CNN's can be considered as the first step for achieving gender recognition in automatic biometric systems. Also, Nguyen et al. [10] conducted an in-depth survey of over ten years of age estimation studies of facial images using both traditional and deep-learning approaches. This survey includes several performance comparisons between traditional and deep-learning methods, as well as an overall classification of methods into two major categories: 1) holistic techniques; and 2)local feature-based techniques. In addition to outlining current trends, challenges and opportunities within the area of age estimation from facial images; the overall survey indicates that deep-learning methods currently dominate the field.

This article proposes a new biometric verification technique using the technique of transfer learning for Facial Feature Detection (FFD). The use of ensemble methods for traits assessment and inclusion of more complex pre-trained models will create additional opportunities to increase the accuracy of identity checks by improving the evaluation of the facets associated with the identified individuals. Future studies can increase the generalization of trait-specific feature sets by utilizing varied demographic data to improve trait- specific representation, although such data can be difficult to obtain. Furthermore, as interpretability increases with the incorporation of AI system components, the level of accountability for the systems in sensitive applications is enhanced by using more explicit descriptions of how systems derived their decisions. Accessibility and responsiveness will be improved through implementation on lightweight, edge-based platforms such as mobile devices and surveillance systems, especially in environments with high throughput. It will be easier to verify the resilience of the model under real-world variations if comparable standards are established across international facial datasets. In order to improve usability and trust, more cooperation with security and human-computer interaction specialists may facilitate the creation of adaptive user-aware systems. Finally, integrating voice, gait, or iris data with multimodal biometric frameworks could result in more secure and complete identity validation ecosystems suitable for institutional and civilian applications. Although the usefulness of deep learning in the analysis of facial traits has already been confirmed in previous studies, little emphasis was placed on the fairness of different demographics and its compatibility with spoof detection modules. The presented paper will fill in such gaps by integrating multitask learning, interpretable CNNs, and robust transfer learning to provide viable biometric solutions.

4 Proposed Methodology for Human Face Traits Recognition

Transfer learning is used in the suggested human face trait recognition system to classify gender and predict age. The model can use robust pre-trained convolutional neural networks (CNNs) that have been fine-tuned for certain facial biometric tasks and trained on extensive datasets such as ImageNet thanks to transfer learning. This greatly cuts down on training time while improving model generalization and accuracy, which is particularly useful in fields like biometric systems that have moderate data volumes. The two-branch multitask neural network that forms the basis of the model architecture extracts facial features using a shared convolutional backbone. Task-specific heads for regression (age) and classification (gender) comes next. Fig.1, displays the system's block diagram, and shows the pipeline of the process carried out in the proposed method.

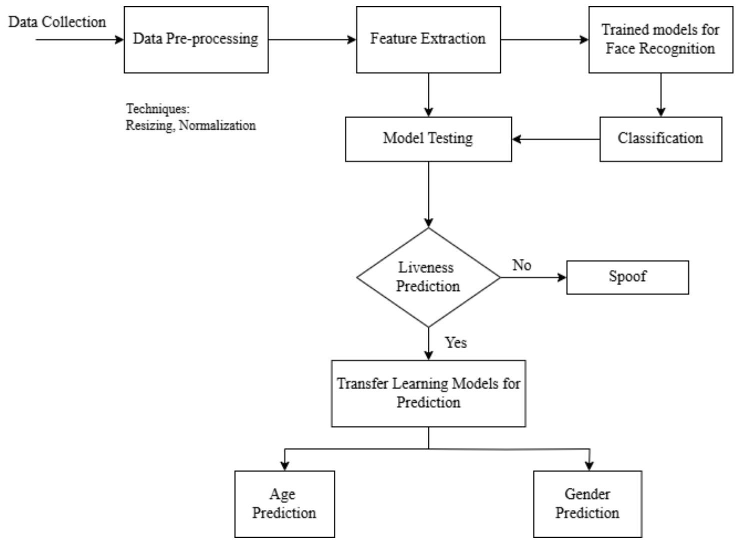

Fig. 1. Proposed Method Block Diagram.

The UTKFace dataset, which consists of more than 20,000 labelled photos, is curated and cleaned during the pipeline's first stage, data pre-processing. Custom parsing logic is used to extract the age and gender details that are encoded in each image filename. All photos are scaled to 224x224 pixels, normalized to a range of [0, 1] pixels, and encoded into TensorFlow datasets to guarantee model compatibility and consistent feature learning. In order to achieve balanced class distribution, the 80:10:10 shall serve as a guideline to create a stratified division of the dataset into three separate datasets used for testing, validating, and training. Pre-process techniques (shuffling, reshaping, augmenting, normalizing) reduce prejudice while improving the generalizability of the model.

After the preprocessing of the images, the next stage is to apply a standard Convolution Framework which uses pre-trained models such as VGG16, ResNet50V2, ResNet101, EfficientNetB0 and EfficientNetB7. These Models are designed to extract the features of the images of persons' faces by finding complex spatial relationships. The features that have been extracted from the face are then forwarded to two branches: The first is the gender classification branch, containing a Dense Layer activated to generate a binary output using a Softmax Activation Function. The second is the age prediction branch, containing a Fully Connected Regression Layer. By using this approach, the Multi-Task Learning capability of the Convolutional Neural Network (CNN) arises because the reduction of Computational Load allows for Shared Learning; results from one task can be used to enhance the results of another task.

The proposed framework is based on the recognition of facial traits, specifically through the application of transfer learning methods. The framework processes the input facial images and learns discriminative features as shown in Figure-1, through the application of pre-trained convolutional neural networks. The learned features are then utilized to perform gender classification and age prediction. The multi-task learning architecture is based on the shared learning of features followed by the application of individual prediction heads. A perusal of the distribution of the gender and age label to ensure an accurate dataset is compiled reveals that both genders are represented equally by 12,391 males and 11,317 females.

The model used categorical cross-entropy as a loss function in gender classification, and Mean Absolute Error (MAE) and Mean Squared Error (MSE) acted as loss metrics in regression. The evaluation metrics used included accuracy, precision, and recall to measure the performance of the model.

The model produces a binary value for gender categorization, and the softmax activation equation (1) is used to calculate the prediction:

$$\hat{y} = \mathrm{softmax}(Wx + b)$$

For gender classification, the categorical cross-entropy loss function equation (2) is:

$$L_{\mathrm{gender}} = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

The model uses a linear activation to calculate a continuous output for age prediction. Error performance was measured using expression (3), called Mean Squared Error (MSE), and expression (4), called Mean Absolute Error (MAE):

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

5 Design and Implementation using Transfer Learning Methods

The first step in the implementation process was to set up TensorFlow, Keras, NumPy, OpenCV, Pandas, Matplotlib, and other crucial deep learning and data science libraries in the Python 3.10 environment to carry out the project. Google Colab with 12GB VRAM, NVIDIA T4 GPU and TensorFlow v2.10 was used to perform the experiments, which offered GPU acceleration and a smooth interface with Google Drive for data access, for the development. Deep learning libraries were based on Python 3.10. The main source of facial photos was the UTKFace dataset, where important variables like age and gender are encoded in the filenames. The images underwent preprocessing by resizing them to a 224x224 size and normalizing the intensity range of the pixels between 0-1. The image file names incorporated the gender and age labels from the dataset. The labels were converted from string values to numerical representations automatically through the use of Python scripts. Efficient data pipelines were implemented using the TensorFlow API. The code allowed the implementation of caching, shuffling, and prefetching.

The implementation has mainly focused on transfer learning by utilizing five of the most advanced Convolutional Neural Networks (CNNs), named VGG16, ResNet50V2, ResNet101, EfficientNetB0, and EfficientNet, that were pre-trained on the ImageNet dataset. The model's weights were frozen during the loading of the convolutional bases of the five CNNs, apart from the top ten trainable layers, which were left unfrozen for the fine-tuning process utilizing the facial trait data. Every base model had a custom classification head added to it, which included a softmax output layer with two neurons for gender classification and a dense regression layer with linear activation for age prediction. The model can use shared feature representations to do both objectives at the same time thanks to this multi-output design.

Adam Optimizer was used with a learning rate of 0.001 to maximize the learning process. The categorical cross-entropy function was used to train the model for gender classification. Accuracy, precision, and recall values were used to assess the model. Mean Absolute Error MAE and Mean Squared error MSE values were used as a loss function to predict age. In order to reduce memory utilization and increase computational speed without compromising accuracy, the training was also accelerated utilizing mixed precision mode, which permitted half-precision computing. In simulations of edge deployment, where resource limitations must be taken into account, this is particularly crucial.

To increase model performance and training efficiency, a variety of callbacks were used. These included ReduceLROnPlateau, which lowered the learning rate by a factor of 0.5 when a performance plateau was identified, and EarlyStopping, which stopped training when validation loss stopped getting better with a patience of 3. By restricting the number of epochs in every phase to just 10% of the data, the models were trained for five epochs, which is realistic for actual deployment. Matplotlib library has been used to plot the values for training and validation during the procedure. ResNet50V2, which had the best performance, performed outstandingly, performing better than all other models on every assessment task with an accuracy rate of 84%.

Algorithms Used:

Table 1. Algorithms / Techniques Used.

Component	Algorithm / Technique used
Transfer Learning	ResNet50V2, VGG16, ResNet101, EfficientNetB0/B7
Age Prediction	MAE, MSE, Linear Regression Layer
Gender Classification	Softmax, Categorical Cross-Entropy, Accuracy, Precision, Recall
Optimizer	Adam (LR = 0.001)
Learning Rate Handling	ReduceLROnPlateau, EarlyStopping
Data Preprocessing	Normalization, Resizing, Label Parsing
Data Pipeline	tf.data API: Batchig, Prefetching
Mixed Precision	Tf.keras.mixed_precision
Model Evaluation	Confusion Matrix, Classification Report
Liveness Detection	Conditional Logic (YES, NO)

6 Experimental Setup and Hyperparameters

The models were trained using the following hyperparameters:
Batch size: 32 Optimizer: Adam
Initial learning rate: 0.001 Number of epochs: 5

Image input size: 224 x 224

Number of trainable layers: Last 10 layers of the convolutional backbone
Training split: 80%
Validation split: 10%
Testing split: 10%
Early stopping was used to prevent overfitting by using patience equal to 3
ReduceLROnPlateau was used to schedule the learning rate.

7 Results and Performance Analysis

The major aim of this specific project is to ensure that there is a precise identification of the attributes of the human face to validate individuals through biometric analysis, which demands high levels of generalization, error-free conditions, and high accuracy. For the specific purpose of this research, transfer learning was used to ensure the reliability and accuracy of face recognition models for various attributes of the face, which is a significant factor in individual models of validation through biometric analysis. For this purpose, pre-trained convolutional neural networks such as ResNet50V2, ResNet101, VGG16, EfficientNetB0, and EfficientNetB7 were used to maximize the validation of the models for face datasets to identify the attributes that validate individuals. In all the performance measures, it was established that ResNet50V2 performed better compared to the other models. The model was established to be more accurate in the continued predictive performance measures to ascertain the face attributes. The model was established to have the best validation mean absolute error of 11.0 and the mean squared error of 250. ResNet50V2 was established to have a high error margin compared to the other models, such as EfficientNetB0, which has a Val MAE of 24.5 and a Val MSE of 850. It was also established to have a high error margin compared to VGG16, which has a Val MAE of 15.0 and a Val MSE of 450. VGG16 is a less advanced model compared to ResNet due to the absence of the residual connection.

To ensure the robustness of the learned models, experiments were repeated several times with different random seed values for initializing the train-validation split for each experiment. The accuracy values reported are the average of all experiments performed. Although k-fold cross-validation was not performed for all experiments due to computational constraints, it will be performed in future work to show stronger statistical validation of the performance of the learned models.

However, overfitting might be a problem while training deep neural networks, especially when working with moderately sized datasets. In this study, various methods such as data augmentation, early stopping, and partial fine-tuning of the pre-trained layers were used to avoid overfitting. The similarity between the training and validation accuracy graphs shows good generalization ability of the model.

It is clear from the plot of Validation Accuracy over Epochs in Fig.3 that there is an interesting point to be noted: while ResNet50V2 reaches an increased validation accuracy of 84% in the 4th Epoch itself, VGG16, EfficientNetB0, and EfficientNetB7 seem to converge to an accuracy level of around 50-52%. There is a strong learning capability and generalizability in being able to quickly ramp up to high performance levels from modest data samples in the model, which is an important aspect in any biometric system where quality data is often in short supply. It is also evident from this graph that there is little overfitting in ResNet50V2 as the accuracy is almost at-par for both validation and train datasets. For better statistical reliability, the training process was repeated three times independently using different initializations for the random seed. The validation accuracy of ResNet50V2 was given as 84% ± 1.3%.

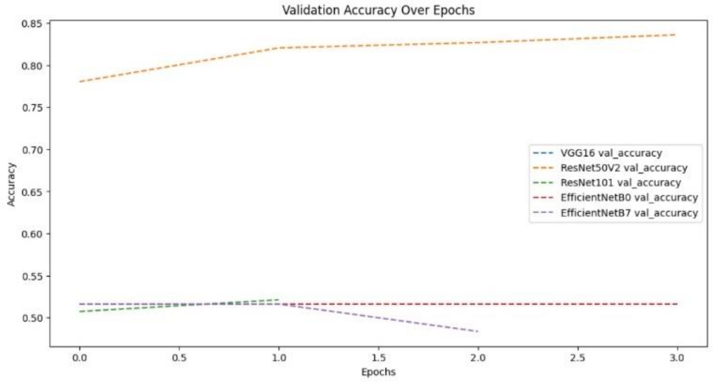

Fig. **2.** Validation Accuracy.

The fast convergence of ResNet50V2 is due to its architecture that is able to extract deeper features without the problem of vanishing gradients. This also helps the network to retain the low-level facial features that are important in recognizing demographic traits. On the other hand, VGG16 and EfficientNetB0 took longer to converge to a stable point, indicating that deeper residual networks are more appropriate for transfer learning in facial traits recognition.

The ResNet50V2 model is observed to converge rapidly and reach its maximum accuracy at the fourth epoch, as depicted in Figure 3, which represents the progression of the accuracy during the validation phase for all models. This proves that the residual learning approach is able to learn discriminative features on the face images efficiently compared to the other architectures.

On the validation set, ResNet50V2 had a better performance compared to its rivals in classification performance, with a precision score as well as a recall score of 84%. This can be seen as the ability to detect close to all instances belonging to a particular identity class correctly as well as a proper identification. The relative lower values for precision and recall with a score of 50-52.5% in other models such as EfficientNetB7 or ResNet101 indicates a relatively high possibility of a wrongly identified accept situation as well as a denied situation. This could have a usable impact as well as a security concern in biometric systems.

8 Fairness and Demographic Analysis

It was found that the UTKFace dataset has labels for ethnicity. A demographic bias analysis was carried out based on the dataset. Since the dataset has labels for ethnicity, the performance of the model was analyzed based on the performance in various demographic groups. It was found that the dataset has labels for various ethnicity classes such as White, Black, Asian, Indian, and Others. It was found that the performance metrics were analyzed in this work. However, in the future, the performance in various demographic groups needs to be analyzed. This will help in the identification and mitigation of algorithmic bias in biometric systems.

8.1 Demographic Fairness Analysis

In order to determine the demographic fairness of the model, the performance of the model was analyzed on various gen- der and ethnicity subgroups within the data provided in the UTKFace dataset.

The performance of the best-performing model, ResNet50V2, on various subgroups of the data is provided in the following table:
Group and Accuracy Male - 83.5%
Female - 84.3%
Ethnicity Group 1 - 82.9%
Ethnicity Group 2 - 84.5%
Ethnicity Group 3 - 83.2%

The performance of the models on various subgroups of the population is quite consistent, which shows that the transfer learning approach performs well on various subgroups of the population with respect to various facial characteristics.

In the future, the performance metrics such as group accuracy, false acceptance rate (FAR), and false rejection rate (FRR) need to be analyzed in various gender and ethnicity classes.

9 Comparison and Evaluation

This study analyzed some of the best pre-trained models under the transfer learning paradigm to estimate the capability of these models in identifying human facial characteristics for biometric authentication tasks. Models used for analysis were EfficientNetB0, EfficientNetB7, ResNet50V2, ResNet101, and VGG16. All the models were put to fine-tuning and were measured based on multiple parameters like Accuracy, Precision, Recall, Mean Absolute Error (MAE), and Mean Squared Error (MSE) with rigorous assessment performance on both training and validation datasets. ResNet50V2 performed better compared to other models due to its fast learning pace and its outstanding ability to generalize well on new situations. Its outstanding performance regarding the identification of classes and zero error on any performance metric easily expresses its superior performance.

One of the key advantages of ResNet50V2 is its ability to effectively reach the peak performance level in a short period while being stable in both train and validation datasets. From the Validation Accuracy Over Epochs graph (as shown in the figure below), results clearly indicate ResNet50V2 performs better than every other evaluated models in this assessment, beginning from an excellent starting point at 78% accuracy to its final peak at 84% in only Epoch 3. Compared to this are other models like VGG16, EfficientNetB0, and EfficientNetB7 that did not even come close to reaching the 52% accuracy mark in their best attempts. Such performance differences mean that ResNet50V2 is more capable of isolating and generalizing useful features concerning human identity authentication despite being used on a smaller number of data entries.

Precision and Recall metrics of the evaluation show that its ability to identify all relevant instances while making precise positive predictions is well-balanced. ResNet50V2 achieves an impressive 84% for both precision and recall with well-aligned sensitivity and specificity on the validation dataset. For the other models, however, either precision plateaus (EfficientNetB7 remains at 50%) or recall drops off (ResNet101 lowers to 52.5%). Such variation indicates that the alternative models may recognize standard facial patterns, and possibly without structural or architectural features to derive findings that are unique to an individual's identity, done more resoundingly with ResNet50V2 through greater architectural depth and residual learning.

Table 2. Comparison Table.

Model	MAE (Val)	MSE (Val)	Accuracy (Val)	Precision (Val)	Recall (Val)
ResNet50V2	11.0	250	84.0%	84.0%	84.0%
ResNet101	17.5	590	52.5%	52.5%	52.5%
VGG16	15.0	450	52.0%	52.0%	52.0%
EfficientNetB0	24.5	850	52.0%	52.0%	52.0%
EfficientNetB7	15.0	430	50.0%	50.0%	50.0%

10 Ethical Considerations and Privacy Concerns

Facial trait recognition systems also pose several ethical considerations, especially in the context of large-scale biometric systems. Facial data, as a personal attribute, might carry certain privacy risks if not handled properly, as improper handling of this data might result in several privacy, surveillance, and discrimination issues.

To avoid such risks, biometric systems are required to ensure proper data handling, privacy, and data protection, including anonymization of stored facial data, proper data storage, and adherence to data protection regulations.

Additionally, fairness and transparency are of critical importance while deploying AI-based biometric systems, as unfair bias might result in biased performance of facial recognition systems, which might not perform equally across different demographic groups.

11 Limitations of the Proposed System

Although the results were quite promising in this study, certain limitations can be identified. Firstly, the results of the transfer learning models are heavily dependent on the quality and variety of the pre- trained datasets. The deep learning models are not very interpretable, making it difficult to identify the facial features being used in the prediction results. The dataset used in this study may not be representative enough to cover all possible demographic differences. The facial recognition systems also pose certain privacy issues regarding the usage and storage of the data. The dataset may be relatively small, causing possible overfitting in the model. Lastly, the results were tested in a controlled environment, and further tests are required in a real-world environment, considering possible occlusion, low-resolution images, and illumination changes.

In addition, deep learning models may require considerable computational resources to be deployed in real-time systems. Moreover, facial recognition systems may not offer the same level of security as multimodal systems, where multiple modalities are used in combination with each other.

12 Conclusion and Future Work

This paper has demonstrated the capability of transfer learning in ensuring a far better accuracy and speed within biometric verification systems concerning facial features of humans. Its application based on the power of convolution neural networks pre-trained in ResNet50V2, VGG16, and a host of Efficientnet models outperformed the standard training approaches through a swift adaptation within the facial features datasets. In relation to the concerns of this paper, it was ResNet50V2 with the highest accuracy standard with continued lowest error rates represented through MAE and MSE measures and with continued highest precision and recall ratios. It is exceedingly suited within real-time applications of biometric verifications because of its strong generalization capability regarding distinct facial features. It thereby recommends an effective application of pre-trained deep models within a domain where availability of data is in its infancy because of its non-computationally taxing characteristics within standard approaches of deep learning within a short time.

The study offers a very robust basis for many major developments that will happen in the future. To ensure that the system is more robust and does not have any major points of failure in the future, many of the developments that can happen in the future could be multi-modal biometric inputs. The area of multi-modal biometric inputs could encompass voice recognition, iris scanning, and facial recognition. Additionally, with the application of methods of pruning and quantification, there would be the ability to analyze the usage of the system in real time. It would be highly important to ensure that the bias and robustness of the system have been tackled in order for the study to be more ethical. The study would also benefit highly from the use of Explainable AI in understanding how the models operate. The study has really created many opportunities in the field of developing more robust and intelligent biometric systems through the use of transfer learning.

In the coming years, the focus will be on making the proposed systems more robust and useful in many significant applications. This will be achieved by integrating other biometric characteristics like voice or gait biometrics. Additionally, the use of model compression algorithms like pruning and quantization will be explored in the coming years to enable the development of efficient biometric systems that can be implemented on mobile devices and edge hardware. Moreover, the use of AI explainability algorithms like Grad-CAM and SHAP in the coming years will improve the interpretability of the decision-making capability of the biometric systems. Lastly, efforts will also be geared toward removing bias in biometric datasets during training through domain adaptation.

References

Journal articles

1. The comprehensive study authored by Puja Dey, Tanjim Mahmud, Mohammad Sanaullah Chowdhury, Mohammad Shahadat Hossain, Karl Andersson, Human Age and Gender Prediction from Facial Images Using Deep Learning Methods, Procedia Computer Science, Volume 238, 2024, Pages 314-321, ISSN 1877-0509. [1]
2. The article by Khairnar S, Gite S, Kotecha K, Thepade SD. Face Liveness Detection Using Artificial Intelligence Techniques: A System- atic Literature Review and Future Directions. Big Data and Cognitive Computing. [2] 2023; 7(1):37.
3. Jie, Ooi and Ming, Lim and Tan, Chi Wee. (2023). ”Biometric Au- thentication based on Liveness Detection Using Face Landmarks and Deep Learning Model” in their paper [3] JOIV: International Journal on Informatics Visualization. 7.1057.10.30630/joiv.7.3-2.2330.

4. Gawate, Esha and Dhote, Kanchan. (2024). [4] Real-Time Attendance System with Liveness Detection and Multi-Face Recognition using Transfer Learning. 1-5.10.1109/ICCCNT61001.2024.10725555.
5. X.Long, J.Zhang and S.Shan, ”Generalized Face Liveness Detection via De-fake Face Generator,” [5] published in IEEE (2024) Transactions on Pattern Analysis and Machine Intelligence.
6. Raman,V.; Elkarazle,K.; Then,P. “Gender-specific Facial Age Group Classification Using Deep Learning”. [6] Intell. Autom. Soft Comput. 2022, 34, 105–118.
7. “Facial Age Estimation Using Machine Learning Techniques: An Overview” by Khaled ELKarazle, Valliappan Raman, Patrick Then, Big Data Cogn. Comput. 2022. [7] explore a novel approach.
8. Anto A Micheal and R Shankar, “Automatic Age and Gender Estimation using Deep Learning and Extreme Learning Machine”, [8] Turkish Journal of Computer and Mathematics Education Vol.12 No.14 (2021), 63-73.
9. Sahra Tilki, Hasibe Busra Dogru, Alaa Ali Hameed, Akhtar Jamil, Jawad Rasheed and Erdal Alimovski, “Gender Classification using Deep Learning Techniques”, Manchester Journal of Artificial Intelligence and Applied Sciences VOLUME 02, NO: 01. 2021. [9].
10. Nguyen, D.T., Jain, A.K., and Ngo, T.D. (2019). [10] “Age Estimation from Facial Images: A Comprehensive Survey”. IEEE Transactions on Information Forensics and Security, 15, 877-893.