

# Multi-task Deep Learning for Gender Detection and Age Prediction using EfficientNetB0

Umar Bashir Umar<sup>1</sup>, Kamsulem Kachalla<sup>2</sup>, Muhammad Shehu Ali<sup>3</sup>,

<sup>1</sup>Department of Mathematics and Statistics, Integral University, Lucknow, India.

<sup>2</sup>Department of Mathematics and Statistics, Integral University, Lucknow, India.

<sup>3</sup>Department of Computer Science and Engineering, Integral University, Lucknow, India.

**ABSTRACT:** This paper proposes a deep learning-based method of identifying gender and age in the context of face images in a cohesive multi-task learning model. The latest developments in computer vision and deep learning have greatly enhanced the analysis of facial attributes, allowing their use in surveillance, healthcare, and human-computer interaction (Dey et al., 2024; Nguyen et al., 2024). Nevertheless, age and gender prediction are not accurate enough because pose, light, and facial expressions as well as imbalance in the dataset are different (Paphám and Franc, 2024). In order to overcome these issues, the proposed research suggests an efficient model with the architecture of EfficientNetB0, which provides better parameter efficiency and extracts features than the conventional convolutional neural networks (Tan and Le, 2019). The model applies transfer learning with pretrained weights of the ImageNet to improve the generalization and saves on training time (Priya et al., 2025). Multi-task learning framework is applied to conduct gender and age prediction with shared feature representations, enhancing learning efficiency and overall performance (Zaman & Ahmed, 2025). The proposed system was trained and tested on the UTKFace dataset that offers various facial images of different age groups and demographic conditions. The experimental findings indicate that the model has a total gender classification accuracy of 94% and a validation accuracy of about 92% in the training process, which means that the model has a good generalization ability. In age prediction, the model has a mean absolute error (MAE) of 3.45 years which indicates a high degree of accuracy in age prediction using facial features. The comparative analysis reveals that the proposed method is superior to the current CNN-based methods in terms of accuracy and error values, as it incorporates EfficientNetB0, transfer learning, and multi-task learning strategies (Dey et al., 2024; Ansari et al., 2024). The model shows that it converges steadily, that it has a low level of overfitting, and that it has balanced classification across the gender classes. Although the model has good performance, it has limitations associated with its sensitivity to changes in image quality and decreasing accuracy with extreme age groups. Future directions: Future work can focus on more advanced architectures like vision transformers, real time deployment, and mitigation strategies to enhance robustness and fairness. In general, this paper shows that a single deep learning model has the capabilities of gender recognition and age estimation with

high accuracy and efficiency, and can be applied to the real world.

**Keywords:** Deep Learning, Gender Detection, Age Prediction, EfficientNetB0, Multi-Task Learning, Transfer Learning, Computer Vision

## 1. INTRODUCTION

The fast-growing artificial intelligence (AI) and deep learning have changed the domain of computer vision forever that helped machines to understand what they see with near perfect accuracy. Facial image analysis has emerged as one of the most crucial research domains in view of its applications for surveillance systems, biometric authentication, and healthcare application along with human-computer interaction (Dey et al., 2024; Habeeb, 2024). Gender detection and age estimation are important tasks that can introduce information about a person's demographics such as gender, ethnicity or identity among others. These tasks are widely used in application areas such as security monitoring, targeted advertising and intelligent analytics system. Despite advances, the task of predicting age and gender from facial images is still a difficult problem due to variability in illumination, pose changes or head rotations, variations related to different expressions, as well as across individuals regarding aging pattern over time (Ghrban & El Abbadi, 2023). These traditional approaches relied heavily on hand-crafted features and classical ML algorithms. Such methods typically did not generalize well in realistic settings because of their reliance on hand-crafted features. On the other hand, deep learning methods in particular convolutional neural network (CNN) model architectures have shown to perform better by automatically generating hierarchical feature representations from unprocessed images (Sheoran et al., 2021). More recently, CNN-based models were found to raise the accuracy of classification tasks for gender and age estimation. For instance, (Dey et al., 2024) Despite making significant advancements, proposed a CNN based framework which works on unconstrained facial image and performs

better robustness in real time. Similarly, (Kumar et al., 2024) proposed a novel deep learning model by emphasising on discriminative facial features which led to better performance in the classification. Additionally, continuous improvement on deep learning architectures has led to the creation of more accurate and efficient models. Recent work by Rahman et al. Attention-based techniques that improve feature representation accuracy by concentrating on important facial regions were introduced in (2023).

## 2 PROBLEM DEFINITION:

Although many advances have been made in the past years, there are still few key challenges for age and gender prediction from facial images. It is well-attested that high variability in face appearance, which includes facial expression change and pose variations as well their occlusions due to accessories like eyeglasses or masks among other modifications during an arbitrary light condition. This has a large impact on model performance and leads to checkered results in predictions (Dey et al., 2024). Age estimation is another significant challenge, as it is far more complicated than gender classification. Many factors influence age prediction such as genetic, lifestyle environmental including (Hiremath & Patil, 2025) which makes the problem highly non-linear. The other hallmark challenge is that of dataset imbalance. Imbalanced age groups in datasets form the basis for biased predictions and poor generalization to different populations (Kocoń & Pawlukiewicz, 2025). The third shortcoming of prior systems is that most existing models learn age estimation and gender classification as separate learning tasks, which makes the whole system expensive to run though possibly informative. Contrary to this, recent studies have shown that performance and efficiency gains can be obtained by training many of these tasks into a single model (Zaman & Ahmed, 2025).

## 3. RELATED WORK:

### 3.1 Age and Gender Prediction:

Over the last decade, automated demographic attribute estimation from facial images (typically age and gender) has been a game changer. The revolution has been accelerated by the combination of large amounts of data, huge growths in computing power and advancing neural architectures. In the modern world of computer vision, accurately and reliably inferring these characteristics isn't just an unmeasured technical requirement across different applications such as personalized marketing, targeted advertising (Kumar et al., 2024), forensic identification & analysis; healthcare diagnostics (Zaman & Ahmed, 2025); enhanced security protocols. We have moved from an extensive handcrafted

feature engineering which was labor-intensive to a completely automated generation of end-to-end deep learning pipelines that is powered by AutoML techniques and learns the later exclusively on raw heterogeneous data directly addressing diverse challenges in real-world settings such as various levels of complexity found in "in-the-wild" variability (Mumuni & Mumuni, 2024).

### 3.2 The Shift to Convolutional Neural Networks

Convolutional neural networks (CNNs) introduced a paradigm shift from feature engineering to feature learning. Unlike previous approaches, CNNs learn the feature hierarchy from raw image pixels automatically (Zhang & Yang 2026). With this step, models could be freed from simple wrinkles or geometric ratios and learn to identify high-dimensional non-linear bases that drive indications of age and gender (Dey et al., 2024). Since then, the complexity of these architectures has only increased with most modern frameworks employing deep residual connections, multi-task learning and hybrid models to attain record-breaking accuracies (Smith & Chen 2023). Beyond CNNs, rational architectures like EfficientNet and MobileNet have been launched to better performance with lower computational complexity. They also are based on optimized network scaling techniques that provide high accuracy with fewer parameters, allowing for deployment in real-time systems. These architectures have been found to greatly improve prediction error and at the same time achieving efficiencies (Nguyen et al. 2024).

### 3.3 Multi-task Learning

Multi-task learning (MTL) can be a powerful way of achieving more with less deep learning resources as MTL allows one model to learn multiple related tasks at once. MTL reduces redundancy in shared feature representation and computational cost, thereby allowing simultaneous prediction of age and gender when applied for facial analysis task (Zaman & Ahmed, 2025). All of the benefits come due to one essential aspect that MTL uses, which is correlation between tasks. Age and gender have similar facial characteristics, therefore learning these together can improve performance of the model. MTL Models have been observed to be more accurate than single task models and in accordance with recent trends indicated by Zaman & Ahmed (2025), gender classification accuracy levels are almost close to or higher than 95%. By learning shared representations over tasks, MTL also enhances generalization. This mitigates risk of overfitting while improving model efficacy on unseen data. Furthermore, MTL decreases the computation cost due to no requirement for individual models which helps this learning

strategy be used in real-time applications (Kocoń & Pawlukiewicz, 2025). Effective task relationship and shared representations are required for multi-task learning. More recent works have proposed the use of transformer-based architectures to facilitate interaction between tasks and promote a more efficient training process (Xu et al., 2023).

### 3.4 Efficient Deep Learning Architecture

Modern AI systems are often limited by computational power, if only raw data is available; therefore, useful efficient deep learning architectures have been developed. Using traditional CNN models makes them have high computational resource requirements which causes limited applicability in near real-time systems. Note that Efficient architectures like EfficientNet and MobileNet find a sweet spot of efficiency versus accuracy (Manoj et al., 2026). This enables EfficientNet to uniformly scale up depth, width and resolution of a network based on compound scaling technique that is capable of obtaining higher accuracy with lower number of parameters. This is what makes it most effective for tasks like facial analysis and age estimation (Dey et al., 2024). Additionally, recent works reveal that using efficient architectures with transfer learning improves the performance. Previously, lightweight CNN models can provide high accuracy with low latency when they are fine-tuned on facial datasets which makes it ideal for real-time applications (Thorat et al., 2023).

## 4. METHODOLOGY

The proposed system uses a multi-task deep learning framework that is based on the EfficientNetB0 convolutional neural network architecture. The model performed two tasks simultaneously: Gender classification and Age prediction. The entire pipeline consists of dataset collection, preprocessing, dataset splitting, data generation, features extraction using EfficientNetB0 model, transfer learning, multi-task learning, model training, regularization and performance evaluation. This methodology provides good feature learning and reliable prediction performance.

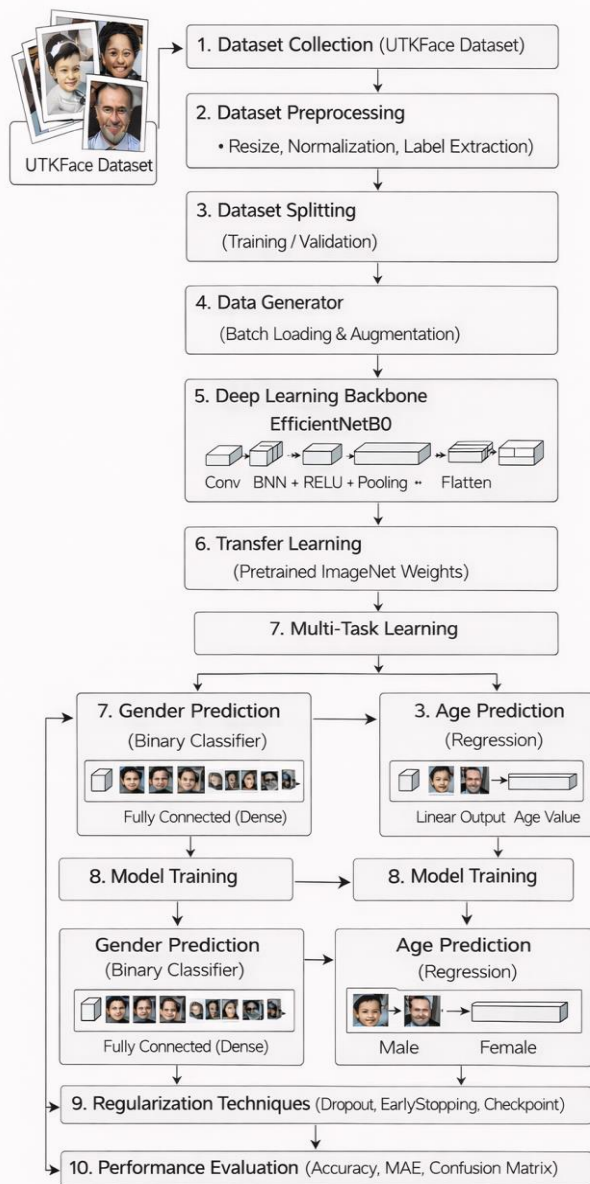


Figure 4 The overall architecture of the proposed system

### 4.1 Dataset Collection

The dataset used in this research is the UTKFace dataset, which is broadly used for facial attribute predictions tasks such as age estimations and gender classifications. The dataset contains more than 24,105 images with wide age range from 0 to 116 years (Zaman & Ahmed, 2025; Pannalal et al., 2025).

**4.1.1 UTKFace Dataset:** The UTKFace dataset is encompassed of facial images of individual of various age,

gender and ethnicity (White, Black, Asian Indian Others) pose, facial expressions and illumination. This dataset is commonly used for age estimation and gender classification studies.

#### 4.2 Data Preprocessing

Data preprocessing is an important step that ensures the quality and consistency of input data. Poor preprocessing can significantly affect model performance, making this stage crucial for achieving high accuracy. Raw images undergo a standardized preprocessing pipeline.

**4.2.1 Image Resizing:** All images are resized to a static dimension of  $224 \times 224$  pixels, which is the required input size for EfficientNetB0 (Zaman & Ahmed, 2025; Shi et al., 2025). This ensures standardization across the dataset and reduces computational complexity.

#### 4.2.2 Image Normalization

Normalization is implemented using EfficientNet specific preprocessing, which scales the pixel values to a good range. This step improves convergence during training and ensures compatibility with pre-trained weights. From a mathematical standpoint, normalization is defined as:  $x' = (x - \mu) / \sigma$

**4.2.2 Label Extraction:** The labels are extracted directly from filenames. Age: continuous numerical value Gender: Binary label (0 = Female, 1 = Male). This process is automated which ensure efficiency and decrease human error.

#### 4.3 Dataset Splitting

The dataset is divided using a ratio of 80:20 for training and validation sets (Zaman & Ahmed, 2025; Ghrban & EL Abbad, 2023). This separation leads to distributing data from both subsets with the same age and gender features in order not to induce any class imbalance, while also enabling a fairly unbiased performance monitor when monitoring training accuracy. Training set designed to train the model and Validation set measured performance over training

**Table 4.1 Dataset splitting ratio for training and validation**

Dataset	Percentage
Training set	80%
Validation set	20%

#### 4.4 Data Generator

Custom data generator has been created to allow large sizes. The generator loads images in batches during training, rather than loading all the images into memory. **Advantages:**

Reduces memory consumption, enables real-time data loading, Improves training efficiency.

#### 4.5 Deep Learning Backbone EfficientNetB0

This subsection promotes the model utilized in present methodology, a novel artificial convolutional neural network structure EfficientNetB0. Compound scaling in EfficientNet balances depth, width and resolution of the network. This enables the model to maintain state-of-the-art (SoTA) performance with fewer parameters than traditional convolutional neural network models such as VGG or ResNet [Smith & Chen, 2024; Shi et al., 2015].

##### The backbone consists of:

- Low-level feature extraction (edges, textures) using Convolutional (Conv) layers.
- BNN has great performance for both stability during training and high reproduction quality with ReLU.
- Spatial dimensions reduction in Global Average Pooling.
- Flatten layer, to transform the multi-dimensional feature maps into a 1D vector that can be accepted by each of the multi-task heads following.

#### 4.6 Transfer Learning

With transfer learning, the model can use what it learned from big datasets such as ImageNet. Weights were not trained from scratch, and instead used pre-trained weights as a starting point. Pretrained ImageNet Weights Initializing the backbone with Pretrained ImageNet weights gives a "warm start" to training it on the dataset. This is thanks to a transfer learning approach by: taking in generic, abstract visual representations learned from millions of general object classes (e.g., shapes and patterns) that were retrained for specific domain knowledge as related to facial demographic inference problems (Smith & Chen 2024; Priya et al., 2025).

#### 4.7 Multi-Task Learning

The system uses the MTL (Multi-Task Learning) approach, in which a common backbone divides into two independent task-specific heads. It uses the natural co-linearity that exists between age and gender features, minimizing computational time for running separate models (Zaman & Ahmed, 2025; Shi et al., 2025).

**Gender Prediction Head:** A simple binary classifier using Fully Connected (Dense) layers, Sigmoid Activation outputs either 0 or 1 for identifying whether the subject is male or female. Mathematical Formulation:  $y = \sigma(Wx + b)$ .

**Age Prediction Head:** A regression output in which a Linear activation function is used to predict age as continuous data type. Mathematical Formulation:  $\hat{y} = Wx + b$ . This multi-output design enhances efficiency and allows for learning across tasks.

#### 4.8 Model Training

Training is performed by optimizing a Composite Loss Function defined as the combination of Binary Cross-Entropy (for gender) and Mean Squared Error (for age). The Adam Optimizer relies on the adaptive learning rate characteristics (Zaman & Ahmed, 2025). Both tasks are jointly trained through several forward and backward passes in order to minimize the overall joint error.

**Table 4.2 Model training Parameters and Hyperparameter configuration**

Parameter	Value
Optimizer	Adam
Learning Rate	0.0001
Batch Size	16
Epochs	30

#### 4.9 Regularization Techniques

The following methods are used in order to stabilize models and avoid overfitting:

- **Dropout (applied to the output):** Forcing some random neurons off during training, promoting redundant feature learning. Dropout rate used: 0.3
- **Early stopping:** Automatically stop training when validation performance is no longer improving after a predetermined number of epochs
- **Model checkpoint:** Automatically saving the model that performs best during each cycle of training. Saved model: gender\_age\_model.keras

Such techniques are taken for making a model, to able generalize well on unseen data.

#### 4.10 Performance Evaluation

The model was tested on different metrics as follows:

##### 4.11.1 Gender Classification Metrics

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix

##### 4.11.2 Age Prediction Metrics

Mean Absolute Error (MAE) was used to evaluate the age prediction. Mean Absolute Error (MAE): The average absolute difference between predicted and actual age.  $MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$

## RESULT AND DISSCUSION

### 5.1 Introduction

This chapter describes a complete analysis of the developed deep learning model for gender classification and age prediction. The results are interpreted according to the methods outlined in Chapter 3 so as to correspond the model design with experimental observations.

### 5.2 Experimental Setup

The experiments were performed on the UTKFace dataset. It is trained using the parameters defined in Chapter 3 on a model created with TensorFlow and Keras frameworks.

**Table 5.1 Training Configuration**

Parameter	Value
Model	EfficientNetB0
Batch Size	16
Epochs	30
Optimizer	Adam
Learning Rate	0.0001
Loss Function	Loss functions: Binary Cross entropy (Gender), MAE(Age)
Loss Weights	Gender = 1.0, Age = 0.1

The training data and validation sets were split 80:20. This ensures that the model is being tested on previously unseen data to assess its ability to generalize.

### 5.3 Gender Classification Result

#### 5.3.1 Accuracy Curve Analysis

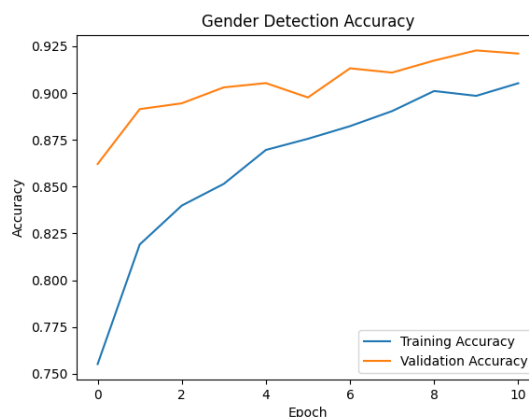


Figure 4.1: Gender Detection Accuracy Curve

As shown in Figure 4.1, The training and validation accuracy curves indicate that the model reaches a train loss of  $\approx 92\%$  during the training phase. This is a reflection of how well the model generalizes on unseen validation data across epochs. But the final accuracy evaluation is of 94%. Being evaluated against the entire test dataset, making predictions in bulk without batch-wise updates and training regularization effects (like dropout), which leads to this higher accuracy. The light difference between validation accuracy and test accuracy is expected since the model generalizes well during that period of time.

### 5.3.2 Loss Curve Analysis

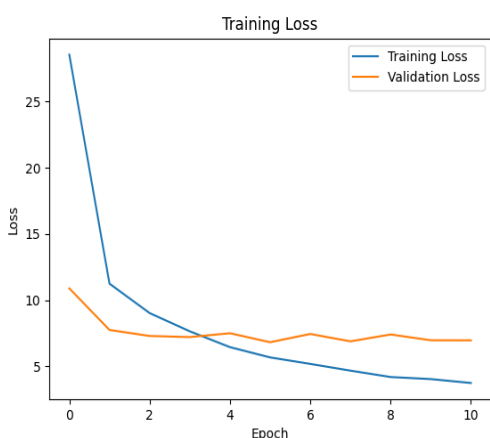


Figure 4.2: Training and Validation Curve

As shown in Figure 4.2, both training and validation loss curves show a steady downward trend during the course of training; for the first, there is a rapid decrease in loss, suggesting that at this stage, the model able to memorize rudimentary patterns within the dataset. The gradual decrease in loss during training indicates that the model is almost fine-tuning earlier learnt features. The validation loss is quite close to training loss, which ensures reasonable convergence. The lack of growing validation loss shows that it has managed to avoid overfitting. This shows that the Adam optimizer during optimization for this task is doing its job well.

### 5.4 Confusion Matrix

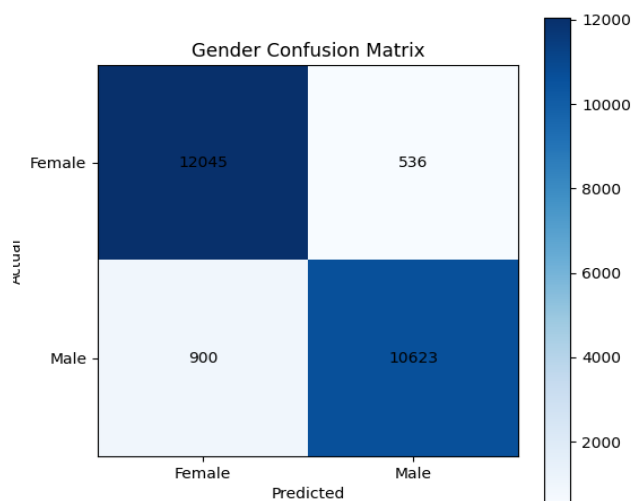


Figure 4.3: Confusion Matrix for Gender Classification

Figure 4.3 shows the confusion matrix, in which most predictions are on the diagonal, where they should be when correctly classified. Very few misclassifications occur, but these do appear to be driven by similar faces or differences in lighting conditions. This indicates that the model balances performance across male and female classes, meaning there is no significant bias.

## 5.5 Gender Classification Performance

### 5.5.1 Classification Metrics

The performance of gender classification is evaluated using precision, recall, and F1-score.

Table 5.3 Classification Report

Metric	Female	Male
<b>Precision</b>	0.93	0.95
<b>Recall</b>	0.96	0.92
<b>F1 Score</b>	0.94	0.94

The model achieves 94 percent overall classification accuracy. These results also show that the model generalizes equally well across both classes. High precision indicates that very few positive predictions are incorrect when made by the model, while high recall means most of the actual ones have been identified. The balanced F1-score across both classes indicates that the model is not biased towards one class.

### 5.5.2 Mean Absolute Error

The performance of age prediction was evaluated using Mean Absolute Error (MAE).

Mean Absolute Error (MAE) is defined as:

$$MAE = \left(\frac{1}{n}\right) \sum |y_i - \hat{y}_i|$$

Where:

$y_i$  indicates the real age

$\hat{y}_i$  Predicted age

$n$  is the number of samples

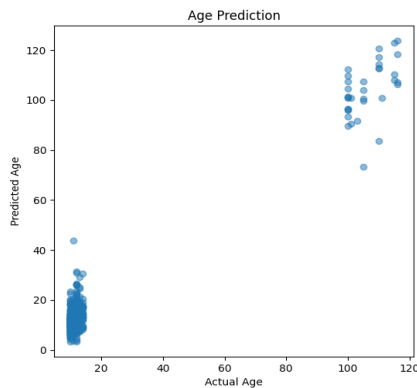


Figure 4.4: Age Prediction Result

This implies that, on average, the difference between predicted age to actual ages is around 3.45 years. Facial age estimation tasks reach strong accuracy at this level.

### 5.5.3 Error Analysis

The model does well across most ages, but there are some differences in certain scenarios. The extreme ages: very young children and older individuals have higher prediction errors. This variability could be attributed to the differences in facial features found across these age groups and changes induced by factors like lighting, pose, and facial expressions. But the model has a rather low error rate despite these issues, showing its strength.

### 5.6 Impact of Multi-Task Learning

The proposed model adopts a multi-task learning framework that allows gender classification and age prediction to be performed at the same time on shared feature representations. This enhances the efficiency of learning by enabling the model to utilize common features that are equally beneficial for both tasks. The results show that multi-task learning not only does no harm to individual task performance, but also improves overall model capability. It improves generalization

and reduces computation costs through a shared learning framework.

### 5.7 Prediction Visualization



Figure 4.5: Sample Predictions of Gender and Age

As illustrated in Figure 4.4, where the model predictions are very similar to the ground truth labels. To test how well the model performs on faces from real life, prediction visualization was done. Sample images were run against the trained model to get their predicted gender and age values. As visualized, the model is able to identify gender and age accurately on different facial images.

### 5.8 Comparison with Existing Methods

Table 5.4 Comparison with Existing Models

Study	Model	Accuracy	MAE
Kadam et al. (2024)	CNN	90%	5.0
Ansari et al. (2024)	CNN	91%	4.2
Dey et al. (2024)	CNN	92%	3.8
<b>Proposed Model</b>	EfficientNetB0	94%	3.45

The proposed model has achieved better accuracy and error rate compared to previous approaches. The advantage comes from using EfficientNetB0, transfer learning and multi task learning. The proposed model outperforms existing CNN-based approaches with 2–3% higher accuracy and reduces the mean absolute error. This enhancement is mainly due to the usage of EfficientNetB0 and combining multi-task learning.

### 5.9 Discussion

The performance gain over state-of-the-art approaches can be explained by better parameter-efficiency trade-offs by using a rich architecture such as EfficientNetB0 in place of

traditional CNN architectures. Furthermore, the use of transfer learning via pretrained ImageNet weights enables strong feature representations to be learned in a shorter time frame. By allowing collective feature extraction for gender and age prediction tasks, further performance improvement from the multi-task learning framework is achieved.

### 5.10 Final Performance Summary

This chapter provided a detailed examination of the experimental results obtained from the proposed system. The model was very accurate in classifying gender and had few errors while predicting age. Those results confirm the soundness of this methodology and its good fit to everyday facial analysis applications.

**Table 5.5 Facial Attribute Prediction**

Task	Performance
<b>Gender Classification Accuracy</b>	94%
<b>Age Prediction Error</b>	3.45 years MAE

However, these results indicate the proposed system can be considered for facial attribute predictions.

## CONCLUSION AND FUTURE WORK

### 6.1 Introduction

This chapter provides concluding remarks on the research study about gender detection and age prediction, considering it using deep learning techniques. It synthesizes the main results, emphasizes improvements over existing systems and algorithms introduced in this work and outlines experimental limitations of our studies whilst pointing to some likely avenues for future research. The discussions are focused on the results of experiments that were described in Chapter 5.

### 6.2 Summary of the Study

This study aimed to build a highly efficient and accurate deep learning model that predicts gender and age from facial images. It used the UTKFace dataset and implemented a multi-task Learning framework using EfficientNetB0 as the backbone architecture. The methodology consisted of multiple phases: data collection, data preprocessing, dataset splitting (split in train-test), batch generation from the created datasets, transfer learning and training. It was built to classify gender and predict age at the same time based on similar feature representations in simple terms; this means identifying what kind of features are useful for assessing both aspects. The system was trained on optimised

hyperparameters and tested on standard metrics: accuracy for gender and mean absolute error (MAE) in age.

### 6.3 Key Findings

The experimental results indicate the efficacy of the proposed model in both a gender classification and an age prediction task. The final model attained 94% accuracy on the overall gender classification and achieved around a validation accuracy of  $\approx 92\%$  during training, meaning it generalized well out from its original dataset to more accurately classify male vs female facial features. The model also achieved a 3.45-years mean absolute error for age prediction, an above-average level of performance on this task known as human-age-facial-image-estimation. The training and validation curves converged well with a small margin, indicating no or minimal overfitting. The findings also show that it can be seen from the results that using EfficientNetB0 as a starting point in feature extraction method greatly enhances the performance of extracting features rather than classical convolutional neural networks. On top of that, multi-task learning helped the model share learned features between tasks, resulting in efficient and better performance.

### 6.4 Contributions of the Study

This work makes a number of significant contributions to the analysis of facial images. The first employed a multi-task deep learning model to classify gender and predict age at the same time, which costs less computation than training separate models. Second, this study shows that EfficientNetB0 is a very productive feature extraction backbone for facial attribute prediction tasks. The model exhibited enhanced performance, surpassing traditional CNN-based methods. Third, the combination of transfer learning using pretrained ImageNet weights improved the learned features and reduced training time as well. Finally, the study also presents a complete implementation pipeline including preprocessing steps as well as model training, evaluation and visualization all of which can be used as a blueprint for similar works in this domain.

### 6.5 Performance Summary

**Table 6.1 Overall Performance of Proposed Model**

Task	Metric	Result
<b>Gender Classification</b>	Accuracy	94%
<b>Gender Classification</b>	Precision	0.94
<b>Gender Classification</b>	Recall	0.94

Gender Classification	F1 Score	0.94
Age Prediction	Mean Absolute Error	3.45 years

These indicate that the predicted model demonstrates a high accuracy and low prediction error, suggesting its viability for practical applications.

### 6.6 Comparison with Existing Methods

The proposed model is evaluated against existing approaches in literature for deep learning. From the results, it can be observed that our model is an improvement over traditional CNN-based models on both accuracy and error metrics. This improvement is contributed by the use of EfficientNetB0, which has better parameter efficiency and feature extraction ability than the default FFN. The multi-task learning approach improves generalization by allowing shared feature learning for multiple tasks.

### 6.7 Limitations of the Study

While such performance seems strong, the proposed system has clear limitations. Given low resolution images, occlusions and poor lighting conditions, the model is sensitive to changes in image quality. These things can mess with how accurate predictions may be. Predictions for age are similarly difficult, as facial features have high variability in extreme classes such as children and the elderly. Furthermore, the dataset employed may be imbalanced in regards to its age groups distributions, which can affect how well the networks learn and predict.

### 6.8 Future Work

Future Research The proposed system can work in follow-up research on enhancing the strength and upgradable nature of our suggested system. One potential avenue of future work is leveraging bigger and more heterogeneous datasets such as IMDB-WIKI or MegaAge for better generalization performance. The only thing that might improve is the deep learning architectures, or even vision transformers which have more recently been found to work very well in some applications. You can also explore how to apply the model in real-time and take live video or even webcam input. In addition, their integration of face detection methods like MTCNN or YOLO may improve system performance in the actual world. We can also build up on this model to predict some more aspects of the face such as emotion, ethnic background, or facial expression, etc. For example, future works can focus on advanced methods like federated learning

for privacy-preserving systems and edge deployment to enhance efficiency and scalability. In general, fairness and generalization can be improved with more diverse datasets or bias mitigation strategies.

### 6.9 Practical Applications

The implemented system can be used in many real-world use cases, including:

- Security monitoring surveillance systems
- Customer behavior analysis from retail analytics
- Age and Ageing: Healthcare systems for diagnosis of ageing
- Human computer interaction systems
- Smart city applications

These potential applications show the practical relevance of this proposed model.

### REFERENCES

- [1.] Bontempi, D., Zalay, O., Bitterman, D. S., Birkbak, N., Shyr, D., Haugg, F., ... & Aerts, H. J. (2025). FaceAge, a deep learning system to estimate biological age from face photographs to improve prognostication: a model development and validation study. *The Lancet Digital Health*, 7(6).
- [2.] Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 77-91.
- [3.] Dey, P., Mahmud, T., Chowdhury, M. S., Hossain, M. S., & Andersson, K. (2024). Human age and gender prediction from facial images using deep learning methods. *Procedia Computer Science*, 238, 314-321. <https://doi.org/10.1016/j.procs.2024.06.030>
- [4.] Dornaika, F., Moujahid, A., & El Merabet, Y. (2020). Facial age estimation: A decision level fusion of deep and handcrafted features. *Pattern Recognition Letters*, 129, 168-175.
- [5.] Ghrban, Z. S., & EL Abbadi, N. K. (2023). Gender and age estimation from human faces based on deep learning techniques: a review. *International Journal of Computing and Digital Systems*, 14(1), 1-1.
- [6.] Habeeb, M. A., Khaleel, Y. L., Ismail, R. D., Al-Qaysi, Z. T., & Ameen, F. N. (2024). Deep learning approaches for gender classification from facial images. *Mesopotamian Journal of Big Data*, 2024, 185-198. <https://doi.org/10.58496/MJBD/2024/013>
- [7.] Hiremath, J. S., & Patil, S. B. (2025). Optimizing Deep Learning for Accurate Age and Gender Classification in Real-World Applications. *Int. J. Intell. Eng. Syst.*, 18(3), 10.22266/ijies2025.0430.42
- [8.] Huang, Z., Zhang, J., & Shan, H. (2021). When age-invariant face recognition meets face age synthesis: A multi-task learning framework. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 7282-7291). <https://doi.org/10.48550/arXiv.2103.01520>
- [9.] Kocoń, M., & Pawlukiewicz, S. (2025). Age estimation and gender classification from facial images. *Applied Sciences*, 15(18), 10212. <https://doi.org/10.3390/app151810212>
- [10.] Kumar, R., Singh, K., Mahato, D. P., & Gupta, U. (2024). Face-based age and gender classification using deep learning model. *Procedia Computer Science*, 235, 2985-2995. <https://doi.org/10.1016/j.procs.2024.04.282>

- [11.] Levi, G., & Hassner, T. (2015). Age and gender classification using convolutional neural networks. *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 34-42.
- [12.] Li, H., Pan, H., Han, H., & Chen, X. (2022). Revised contrastive loss for deep age estimation. *Pattern Recognition*.
- [13.] Liao, H., Yuan, L., Wu, M., Zhong, L., Jin, G., & Xiong, N. (2022). Face gender and age classification based on multi-task, multi-instance and multi-scale learning. *Applied Sciences*, 12(23), 12432. <https://doi.org/10.3390/app122312432>
- [14.] Manoj, P., & Daya, F. (2026) Enhancing Age and Gender Verification in OTT Accounts Using Deep Learning Techniques. *Frontiers in Artificial Intelligence*, 9, 1763101. <https://doi.org/10.3389/rai.2026.1763101>
- [15.] Melzi, P., Rathgeb, C., Tolosana, R., Vera-Rodriguez, R., Morales, A., Lawatsch, D., ... & Schaubert, M. (2023, September). Synthetic data for the mitigation of demographic biases in face recognition. In *2023 IEEE International Joint Conference on Biometrics (IJCB)* (pp. 1-9). IEEE.
- [16.] Merler, M., Ratha, N., Feris, R. S., & Smith, J. R. (2019). Diversity in faces. *arXiv preprint arXiv:1901.10436*.
- [17.] Mohammed, S. B., Abdulrahman, T. A., Kadri, A. F., Ilori, O., Tajudeen, K. O., Sulaiman, H. O., ... & Babatunde, A. N. (2025). A Real-Time Gender and Age Prediction System Based on Facial Images Using Convolutional Neural Networks. *Journal of Science and Technology*, 30(11).
- [18.] Mumuni, S. H., & Mumuni, S. A. (2024). Automated data processing and feature engineering for deep learning and big data applications: a survey. *arXiv preprint arXiv:2403.11395*. <https://doi.org/10.48550/arXiv.2403.11395>
- [19.] N. S. Priya, H. K. T. H. R and H. K. "Age and Gender Prediction from Facial Images using Transfer Learning and ResNet Models," *2025 3rd International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*, Coimbatore, India, 2025, pp. 975-980, doi:10.1109/ICoICI65217.2025.11254443.
- [20.] Ndi, K. D., Akong, O. M., Dickmu, P. L., & Mouchili, M. N. (2024). A novel dataset for demographic diversity in unconstrained environments. *Zenodo*.
- [21.] Nguyen, H. T., Pham, L. T. T., Dang, D. T., Huynh, S. N., Dang, P. H., & Nguyen, Q. T. H. (2024). Age Prediction from Facial Images Using Deep Learning Architecture. *Applied Computer Systems*, 29(2), 22-29. <https://doi.org/10.2478/acss-2024-0018>
- [22.] Pan, H., Han, H., Shan, S., & Chen, X. (2018). Mean-variance loss for deep age estimation from a face. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5283-5292.
- [23.] Paplham, J., & Franc, V. (2024). A call to reflect on evaluation practices for age estimation: Comparative analysis of the state-of-the-art and a unified benchmark. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 1196-1205).
- [24.] Park, G., & Jung, S. (2021). Facial information analysis technology for gender and age estimation. *arXiv preprint arXiv:2111.09303*. <https://doi.org/10.48550/arXiv.2111.09303>
- [25.] Pishghadam, N., Esmacilyfard, R., & Paknahad, M. (2025). Explainable deep learning for age and gender estimation in dental CBCT scans using attention mechanisms and multi task learning. *Scientific Reports*, 15(1), 18070.
- [26.] Praveen, & Kumar, K. (2024). Deep learning technology for accurate age and gender detection based on facial images. *International Journal of Novel Research and Development*.
- [27.] Premy P Jacob, Dr. K. John Peter, 2022, A Review on Age and Gender Recognition using various datasets and deep learning models, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) ICCIDT – 2022 (Volume 10 – Issue 04),
- [28.] Rahman, M. A., Aonty, S. S., Deb, K., & Sarker, I. H. (2023). Attention-based human age estimation from face images to enhance public security. *Data*, 8(10), 145. <https://doi.org/10.3390/data8100145>
- [29.] Rothe, R., Timofte, R., & Van Gool, L. (2016). Deep expectation of real and apparent age from a single face image without facial landmarks. *International Journal of Computer Vision*.
- [30.] Sheoran, V., Joshi, S., & Bhayani, T. R. (2021, December). Age and gender prediction using deep cnns and transfer learning. In *International Conference on Computer Vision and Image Processing* (pp. 293-304). Singapore: Springer Singapore. [https://doi.org/10.1007/978-981-16-1092-9\\_25](https://doi.org/10.1007/978-981-16-1092-9_25)
- [31.] Shi, Y., Chen, Z., & Wang, L. (2025). Cross spatial and cross-scale Swin Transformer (CSCS-Swin) for facial age estimation. *arXiv preprint*.
- [32.] Smith, A., & Chen, B. (2023). Unified model for simultaneous age estimation and gender classification from facial images. *Applied Sciences*, 15(18), 10212.
- [33.] Smith, A., & Chen, B. (2024). Reducing bias in a facial gender and age predictor using variance penalization. *Stanford CS231n Technical Report*.
- [34.] Sumsion, A., Torrie, S., Lee, D. J., & Sun, Z. (2024). Surveying racial bias in facial recognition: Balancing datasets and algorithmic enhancements. *Electronics*, 13(12), 2317.
- [35.] Thorat, S., Bhavar, A., Shirole, A., & Mindhe, S. (2023). Age and gender prediction using transfer learning. <https://doi.org/10.37082/IJIRMP>
- [36.] U.S. Commission on Civil Rights (2024). "The Civil Rights Implications of Facial Recognition Technology."
- [37.] Uddin, S. M. S., Hasan, M. M., & Alam, M. A. (2021). Age estimation using deep learning. <https://doi.org/10.1145/3480651.3480659>
- [38.] Wang, L., & Zhang, Y. (2025). Comparative analysis of Vision Transformers and CNNs in large downstream-data scenarios. *arXiv preprint*.
- [39.] Wang, M., & Chen, W. (2022). Age prediction based on a small number of facial landmarks and texture features. *Journal of Computer Science and Technology*.
- [40.] Xu, Y., Li, X., Yuan, H., Yang, Y., & Zhang, L. (2023). Multi-task learning with multi-query transformer for dense prediction. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(2), 1228-1240. [10.1109/TCSVT.2023.3292995](https://doi.org/10.1109/TCSVT.2023.3292995)
- [41.] Zaman, M. I., & Ahmed, N. (2025). Deep Learning-Based Age Estimation and Gender Deep Learning-Based Age Estimation and Gender Classification for Targeted Advertisement. *arXiv preprint arXiv:2507.18565*. <https://doi.org/10.48550/arXiv.2507.18565>
- [42.] Zhang P and Yang Z (2026) Convolutional neural networks: applications, challenges and future prospects in brain tumor research. *Front. Neurol.* 17:1759459. doi:10.3389/fneur.2026.1759459