ORIGINAL ARTICLE

Deep convolutional neural networks for age and gender estimation using an imbalanced dataset of human face images

İsmail Akgül[1]

Abstract

Automatic age and gender estimation provides an important information to analyze real-world applications such as human–machine interaction, system access, activity recognition, and consumer profile detection. While it is easy to estimate a person's gender from human facial images, estimating their age is difficult. In such previous challenging studies, traditional convolutional neural network (CNN) methods have been used for age and gender estimation. With the development of deep convolutional neural network (DCNN) architectures, more successful results have been obtained than traditional CNN methods. In this study, two state-of-the-art DCNN models have been developed in the field of artificial intelligence (AI) to make age and gender estimation on an imbalanced dataset of human face images. Firstly, a new model called fast description network (FINet) was developed, which has a parametrically changeable structure. Secondly, the number of parameters has been reduced by using the layer reduction approach in InceptionV3 and NASNetLarge DCNN model structures, and a second model named inception Nasnet fast identify network (INFINet) was developed by concatenating these models and the FINet model as a triple. FINet and INFINet models developed for age and gender estimation were compared with many other state-of-the-art DCNN models in AI. The most successful accuracy results in terms of both age and gender were obtained with the INFINet model (age: 61.22%, gender: 80.95% in the FG-NET dataset, age: 72.00%, gender: 90.50% in the UTKFace dataset). The results obtained in age and gender estimation with the INFINet model are much more effective than other recent state-of-the-art works. In addition, the FINet model, which has a much smaller number of parameters than the compared models, showed a classification performance that can compete with state-of-the-art methods for age and gender estimation.

Keywords Deep convolutional neural networks · Model concatenate · FINet · INFINet · Imbalanced dataset · Age and gender estimation

1 Introduction

Facial age estimation is defined as the automatic estimation of either the biological age or age group of a person (child, adult, elderly, etc.), while facial gender estimation is defined as the classification of a person's gender as male or female based on a facial model. Age and gender estimation has vital importance in applications such as consumer profile estimation, social media advertising, demographic profiling, and customized advertising systems [1]. In addition, age and gender estimation can be useful in many areas such as the identification of people to prevent fraud, theft, or anything like that, controlling access to a system, and human–machine interaction [2]. It also plays an important role in smart applications such as healthcare and marketing intelligence [3].

Automatic face and gender estimation, which has become interesting in recent years, is a challenging research topic [4]. When race and gender differences are included, age estimation from a human face image is quite difficult. The performance of the learning model to be used for age and gender estimation largely depends on the data in the dataset. Human face images include many features such as age, gender, race, emotional expression, and health

✉ İsmail Akgül
iakgul@erzincan.edu.tr

1 Department of Computer Engineering, Faculty of Engineering and Architecture, Erzincan Binali Yıldırım University, Erzincan, Turkey

status. Studies on the age estimation problem using facial images date back to 1994. Since then, many different studies have been carried out. Despite extensive work on the age estimation problem, the obtained results still do not show the accuracy and reliability performance to meet real-life demands [5]. The issue of age and gender estimation has been interesting in recent years, both because of the difficulties in age and gender estimation problems and because it still does not fully meet the demands.

Image processing and computer vision have become among the most popular fields with real-world applications [6] and have been used in many different fields with supervised and unsupervised deep learning approaches in recent years [7–14]. After the advent of computer vision and machine learning, automated computerized age and gender estimation systems have become more popular [15]. With the advent of deep learning, the study of facial systems has completely changed. As deep learning has entered many fields, the techniques based on DCNNs have also become a research point in the field of facial age and gender estimation [16]. Estimating age and gender from human facial images is becoming an exciting business in the field of computer vision. Compared to traditional handcrafted methodologies, CNNs and DCNNs, which have been widely used recently during the categorization task, performed much better in age and gender estimation ([17–20]). Significant advances in DCNN architectures such as parallel and hierarchical feature extraction, multi-tasking and transfer learning, data analysis and predictability, and computational efficiency have been effective in increasing performance.

In the literature, many applications of AI methods can be found to estimate age and gender from human face images. Lee et al. proposed a deep residual learning network model consisting of three deep neural networks for age and gender estimation. Model training was done via the IMDB-WIKI database with images collected (more than 14,000) from the internet. In the proposed model, after the faces in the images in the dataset were detected, the age and gender of the faces were estimated. In the augmented IMDB-WIKI dataset with the proposed model, an accuracy rate of 52.2% for age and 88.5% for gender was achieved [21]. Terhörst et al. proposed a new reliability measure neural network that determines the reliability of model predictions for reliable age and gender estimation from face images using the Adience dataset. In their experiments with age groups consisting of eight classes (0–2, 4–6, 8–13, 15–20, 25–32, 38–43, 48–53, and 60 +) and gender groups consisting of two classes (female and male) in the dataset containing more than 26.5 k images, they found that the proposed method was successful in measuring the estimation reliability. From the images in the Adience dataset,

they achieved an accuracy of 64.3% in estimating age groups and 89.8% in estimating gender [22].

Ozbulak et al. proposed DCNN-based generic AlexNet-like and domain-specific VGG-face CNN models to obtain age and gender classification in the wild. The models were used and fine-tuned with the Adience dataset prepared for age and gender classification in uncontrolled environments. As a result of the analysis, 57.9% accuracy in age classification and 92.0% accuracy in gender classification were obtained [23]. In another study using the Adience dataset, a new model was proposed by fine-tuning with pre-trained CNNs. With the proposed model, 62.26% age classification success was achieved [24]. Duan et al. proposed a hybrid structure for age and gender classification that integrates the synergy of two classifiers, including a CNN and an extreme learning machine (ELM). As a result of the analyses, they made using the MORPH-II and Adience benchmark datasets, they achieved a success rate in age classification at 52.3% and in gender classification at 88.2% [25]. Another study using the Morph dataset proposed a face recognition method that does not change according to age. With the proposed method, 92.8% success was obtained [26].

Rwigema et al., in their study, proposed a new hybrid algorithm consisting of conventional artificial neural networks (C-ANNs) and CNN for age and gender classification using a decision fusion technique. They combined the decisions obtained by the two neural networks with probabilistic decision fusion techniques such as majority voting decision fusion, Naive–Bayes combination decision fusion, and sum rule decision fusion. They divided 2000 different images of individuals obtained from the internet into age groups (1–24, 25–49, 50 years, and over) and gender groups. Using the created dataset, an 86.1% accuracy rate in age classification and a 98.4% accuracy rate in gender classification were achieved [27]. Sharma et al. proposed CNN approach improved for face-based age and gender estimation. UTKFace, IMDB-WIKI, FG-NET, and CACD datasets were used for model evaluations. The proposed model trained on a large-scale dataset, UTKFace (aligned and cropped faces), showed 94.01% accuracy for age estimation and 99.86% accuracy for gender estimation [2]. Additionally, many different applications have been implemented for age and gender estimation from human face images using different datasets [28–33, 46].

In addition, many applications of AI methods can be found in the literature for age estimation from human face images with training made using the FG-NET aging database discussed in this study. Nithyashri and Kulanthaivel, in their study using the FG-NET aging database, developed a system on wavelet transformation (WT) to extract facial features and an artificial neural network (ANN) to classify age groups. As a result of their analysis dividing the age

groups into four groups child (0–12 years), adolescence (13–18 years), adult (19–59 years), and senior adult (60 years and above), they achieved a 94.28% correct classification success by using the distance evaluated between the eye center and the mouth (FPD_3) as the feature point distance and Coif Wavelet [34]. Choobeh, FG-NET aging database was divided into two groups (child and adult) and a one-dimensional feature space was created using the active appearance model (AAM) and then linear discriminant analysis (LDA), and the minimum distance classifier on this space was used. As a result of the analysis made with the image-based method created to separate children from adults, children, and adults, they were recognized with an accuracy rate of up to 89% and 90%, respectively [35]. In another study conducted by dividing the FG-NET aging database into two groups (child and adult), statistical modeling of the face was used to separate children from adults according to their facial images. By applying LDA to the face parameters, useful features were extracted, and the Euclidean distance function was used. With the proposed algorithm, an 85% accuracy rate has been achieved [36]. In another study, Razalli et al. presented a two-stage image-based method to distinguish children from adults. According to the face shape elliptical ratio and facial features angle distribution, the analyzes performed using SVM and multi-DVM classification achieved a success rate of 92% [37].

In addition, as a result of the training with different datasets, many AI methods have been applied in the literature where the FG-NET aging database discussed in this study is used as a test dataset. In a study in which FG-NET aging database was used as a test dataset, age progression/regression was determined by conditional adversarial autoencoder (CAAE). It was determined that CAAE showed superior performance with age groups divided as 0–5, 6–10, 11–15, 16–20, 21–30, 31–40, 41–50, 51–60, 61–70, and 71–80 [38]. In another study, Tyagi and Sood used FG-NET dataset as a test with machine learning techniques and examined age groups. An algorithm supporting adaptive features based on local binary patterns and vector machine classification is proposed. With the proposed model, 56% of classification success was achieved [39]. Kumar et al. proposed a method (ADMM + Gabor + SVM) based on Seg-Net-based architecture and support vector machine with machine learning algorithm for age and gender classification from various face images. As a result of the analyses, they made by dividing the age groups into eight classes 0–4, 5–9, 10–14, 15–19, 20–24, 25–44, 45–55, and 56–70, a success of 92.48% was achieved in age classification with the FG-NET dataset [40].

Age and gender estimation contribute to many real-world applications such as controlling access to a system through human–machine interaction, identification and activity recognition in case of a crime, consumer profile detection in the marketing process, and the development of customized advertising systems. Solving the problems of age and gender estimation from human face images has still not been fully met. Since DCNN models perform better than traditional CNN models, using the DCNN approach for age and gender estimation served as the motivation for this study. In this study, several approaches are presented for age and gender estimation from human face images in an imbalanced dataset. First, a new model called FINet was developed. Then, seven different keras models accepted in the literature were discussed, and these models were trained using weights previously trained with ImageNet. From the Xception, InceptionV3, InceptionResNetV2, MobileNet, MobileNetV2, NASNetMobile, and NASNetLarge models discussed in this study, the number of parameters was reduced by using the layer reduction approach to the InceptionV3 and NASNetLarge model structures, and a new model called INFINet was developed by concatenate these two models with the FINet model. To summarize, the FINet model was developed to modulate the structure parametrically with an understanding of simplicity. The INFINet model was developed to obtain diversity in the feature space by producing different features from state-of-the-art DCNNs with the understanding of cost-efficiency without the need for very big data and to prevent overfitting with controls during the learning process. FINet and INFINet models developed for age and gender estimation on an imbalanced dataset were compared with state-of-the-art DCNN models in the field of AI.

When the studies in the literature were examined, it was seen that many different methods were used to estimate age and gender. In this study, two different models are proposed in addition to the studies in the literature. The novelty, contribution, and importance of this study can be summarized as follows:

- Development of the fast identify network (FINet) model for age and gender estimation.
- Performing the triple concatenation of two fine-tuned AI models with the FINet model
- Proposing the inception Nasnet fast identification network (INFINet) model
- Evaluation of age and gender estimation performance of DCNN models on human face images.
- Training and testing the two proposed models and seven different models in the literature using two different datasets in the same environment and obtaining more successful results with the proposed model.

The rest of the study is organized as follows: Materials and methods related to the study are mentioned in

Chapter 2, results and discussion in Chapter 3, and conclusions in Chapter 4.

2 Materials and methods

2.1 Dataset

In this study, face and gesture recognition research network (FG-NET) aging database has been used for age and gender estimation from human face images [41]. Data have been obtained from different sources as the original FG-NET website no longer provides this publicly available dataset [42, 43]. FG-NET is a dataset consisting of a total of 1002 images in different pixel sizes, which are used for age estimation and face estimation between ages and include 82 people between the ages of 0–69. The age distribution in the dataset is not balanced. Images with individuals at their more recent ages are the ones for which digital images were available. For most individuals, images were collected through scanning photographs in personal collections [44]. In addition, when the dataset is analyzed in terms of gender distribution, there are 431 images of 34 women and 571 images of 48 men [45]. In Fig. 1, sample images of different ages of a person in the FG-NET aging database are given.

Each image in the dataset is named to provide information about the image. For example, the meaning of image "048A00.JPG" is the image of person number 48 at age 0, and the meaning of image "048A54.JPG" is the image of person number 48 (the same person) at age 54.

2.2 Image pre-processing

To realize the training process more successfully, image processing techniques such as face recognition and size adjustment were applied to the images in the dataset. First, face recognition was performed using the haar cascade classifier on all images. Later, the recognized face images were adjusted to $224 \times 224 \times 3$ pixel dimensions. In Fig. 2, sample images obtained as a result of pre-processing are given.

The problem can be approached as a classification task by dividing the age ranges into classes [46]. Therefore, each image was collected in different folders according to

certain age ranges, age groups were created and labeled (For example, people aged 0, 1, and 2 were collected in the folder belonging to the 0–2 age group, and people aged 4, 5, and 6 were collected in the folder belonging to the 4–6 age group). In order to make a comparison with the literature, age groups are divided into eight classes: 0–2, 4–6, 8–12, 15–20, 25–32, 38–43, 48–53, and 60 + . Additionally, gender groups are divided into two classes: female and male. Accordingly, the facial region images obtained were grouped separately in terms of age and gender. In Table 1, the number of images in each class according to age and gender groups is given.

2.3 Work environment and training parameters

With the models used in this study, the Google Colaboratory Pro [47] working environment was used to make age and gender estimation successfully on the imbalanced dataset. All operations in the Colab environment with NVIDIA Tesla K80 graphics processor are coded using the Python programming language.

In this study, state-of-the-art models were preferred to obtain successful classification results in age and gender estimation from the imbalanced dataset consisting of human face images. These models are the Xception [48], InceptionV3 [49], InceptionResNetV2 [50], MobileNet [51], MobileNetV2 [52], NASNetMobile, and NASNetLarge [53]. To carry out the training of the preferred models, the dataset is divided into two datasets 80% train and 20% test, separately according to both age and gender. Model training was evaluated with the training dataset, and model performances were evaluated with the test dataset. As a result of dataset splitting, train-test dataset image numbers obtained for each age and gender group are given in Table 2.

There are many hyperparameters in neural network models that affect performance. Many trials were made to determine the hyperparameters to be used in this study. When one parameter changed, the model performance was observed by keeping the other parameters constant and the best hyperparameter value was determined. Therefore, the train and test of each model were carried out using the hyperparameters given in Table 3. In this study, the epoch 30, mini-batch size 8, optimization algorithm SGD, and learning rate were determined as 0.001. Since the age

Fig. 1 Sample images of different ages of a person in the FG-NET aging database

Fig. 2 Sample images obtained as a result of pre-processing in the FG-NET aging database

Table 1 Number of images in each class according to age and gender groups in the FG-NET dataset

	Age groups								Gender groups	
	0–2	4–6	8–12	15–20	25–32	38–43	48–53	60 +	Female	Male
Number of images	109	123	166	185	89	35	16	8	308	423

Table 2 Number of images of the train-test dataset obtained for each age and gender group in the FG-NET dataset

	Groups	Train (%80)	Test (%20)	Total (%100)
Age	0–2	87	22	109
	4–6	98	25	123
	8–12	133	33	166
	15–20	148	37	185
	25–32	71	18	89
	38–43	28	7	35
	48–53	13	3	16
	60 +	6	2	8
	Total	584	147	731
Gender	Female	246	62	308
	Male	338	85	423
	Total	584	147	731

groups are in eight classes, the activation function was determined as Softmax and the loss function as categorical crossentropy in age estimation. Since the gender groups are in two classes, the activation function was determined as sigmoid and the loss function as binary crossentropy in gender estimation. Experiments were conducted using the same environment and the same training parameters to determine the success of all models used in this study.

2.4 Model structures

In this study, a new model named FINet was developed, and a concatenate model named INFINet was created to make age and gender estimation successfully on the imbalanced dataset.

2.4.1 FINet model

In this study, a new DCNN model named fast identify network (FINet), whose schematic diagram is given in Fig. 3, is developed for age and gender estimation on an imbalanced dataset.

In the FINet model given in Fig. 3, a convolution operation with filter, kernel, stride, and padding values of 32, 3, 2, and "valid" in the initial block, respectively, is applied to the input images, and then, the ReLU activation function is applied. In the next block, a convolution operation with filter, kernel, stride, and padding values of 64, 3, 1, and "same", respectively, and after convolution, the batch normalization and ReLU activation function are applied twice. After these blocks, the max-pooling layer with a pool size of 2 and the dropout layer with a drop rate of 0.15 come. These operations are repeated in the next blocks by changing the parameter values. The filter size is increased by 32 in each convolution layer. Other parameters are increased or decreased within a certain logic in
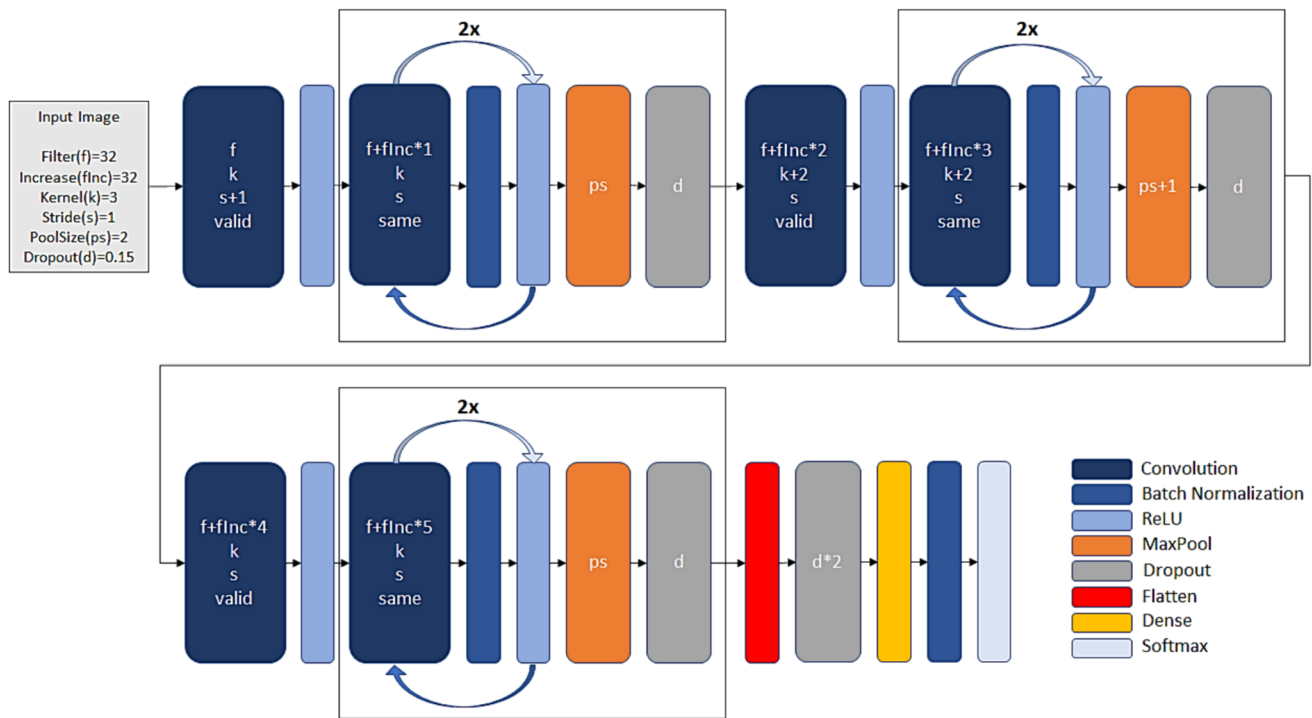
Table 3 Hyperparameters used for train and test of each model

Hiperparameters	Value	
	Age estimation	Gender estimation
Epoch	30	
Mini-batch size	8	
Activation function	Softmax	Sigmoid
Optimizer	SGD	
Learning rate	0.001	
Loss function	Categorical Crossentropy	Binary Crossentropy

Fig. 3 Schematic diagram of the FINet model

each layer. In this way, a new model with a parametrically modifiable structure has been developed.

Each layer used in the FINet model can be summarized as follows:

- Convolution: It was used to obtain an output with new features by shifting a filter on the input data.
- Batch normalization: It was used to make the training of the network faster and more stable by rescaling the input values coming to the neuron in the neural network.
- ReLU (Rectified linear unit): It was used to determine an output value in response to the input value coming to a neuron in the neural network.
- MaxPool: It was used to take the maximum value in the area covered by the filter to reduce the number of parameters of the neural network by reducing the size of the feature map created by the convolution layer.
- Dropout: It was used to prevent overfitting of the neural network by eliminating random nodes.
- Flatten: It was used to transform the two-dimensional arrays obtained from the feature map into a single long vector.
- Dense: It was used to change the size of the vector in a way that is deeply connected to the previous layers.
- Softmax: It was used in the output layer to perform classification in the neural network.

2.4.2 INFINet model

It is seen that DCNN models with attribute combinations perform better than models without attribute combinations [54]. For this purpose, in this study, block cutting was applied to the InceptionV3 and NASNetLarge model structures to create a new and effective model by concatenation of state-of-the-art DCNN models in the field of AI for age and gender estimation on an imbalanced dataset. A series of ablation studies were carried out to evaluate the effects of block cutting on the performance of the models and to show that it is important for the functionality of the cut section. As a result of many trials, the best performance was achieved with the InceptionV3 and NASNetLarge models, so block cutting was applied to these two models. Figure 4 shows InceptionV3 (left) and NASNetLarge (right) model structures obtained as a result of block cutting. In the following sections of the study, the effect of the cut section on the performance of the models, the parameter numbers of the models formed as a result of block cutting, and their comparison with the parameter numbers of other models are given.

After block cutting from the points shown in Fig. 4, the cut InceptionV3 and NASNetLarge models and the FINet model were concatenated. Thus, a new DCNN model named inception Nasnet fast identify network (INFINet) was created. The schematic diagram of the INFINet model is given in Fig. 5.

Fig. 4 InceptionV3 (left) and NASNetLarge (right) model structures as a result of block cutting
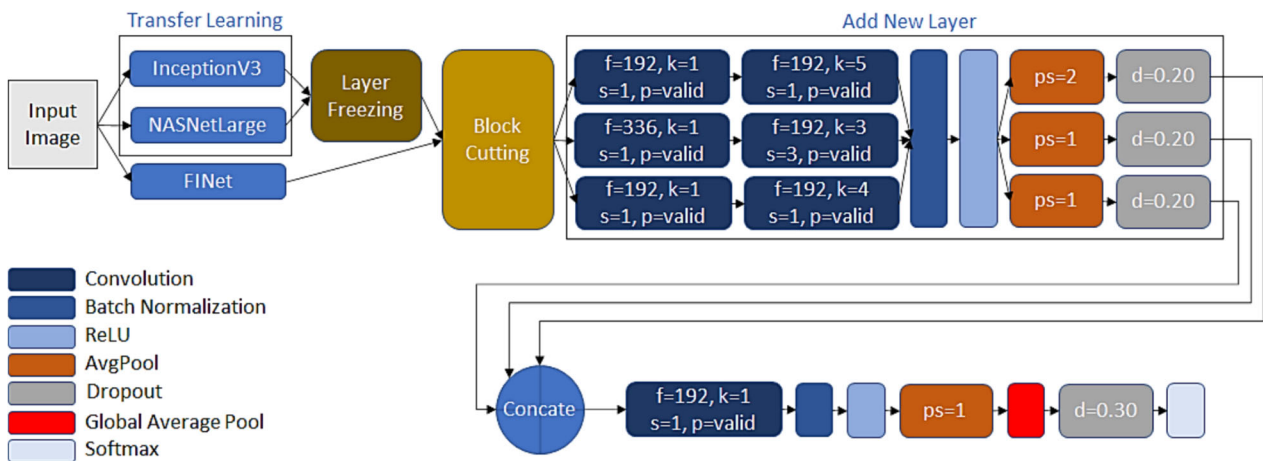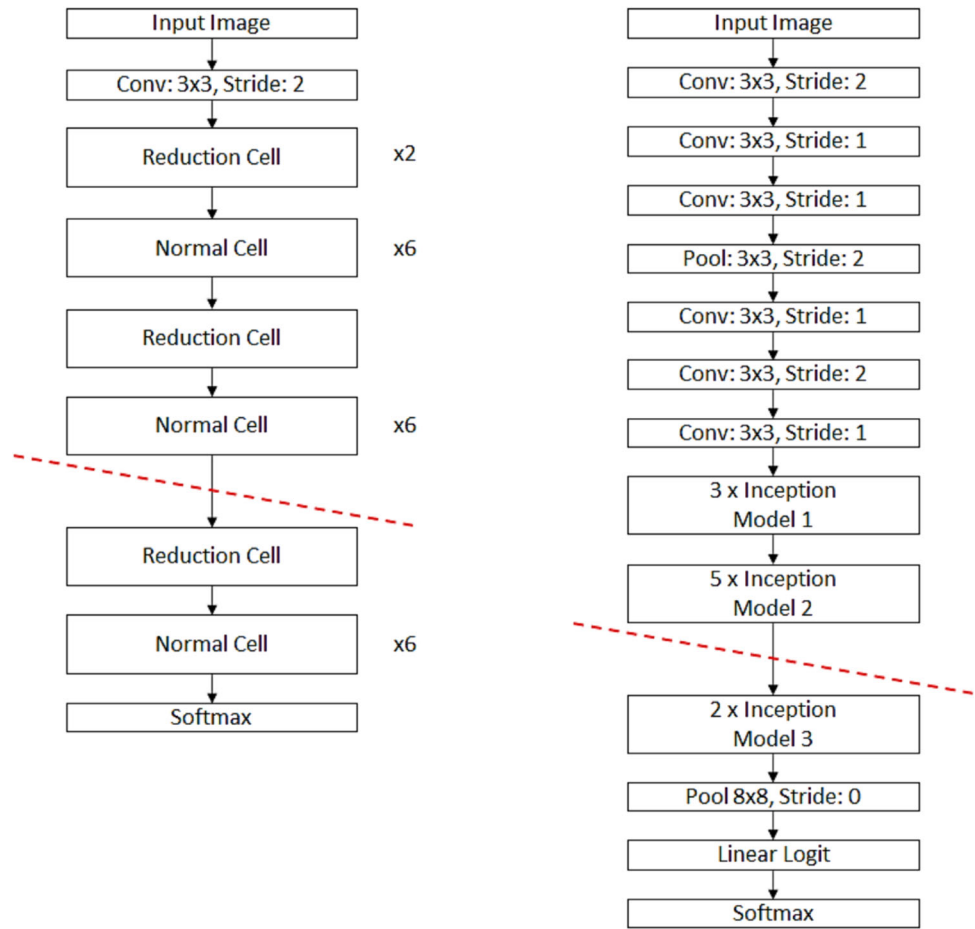

Fig. 5 Schematic diagram of the INFINet model

In the INFINet model given in Fig. 5, firstly, layer freezing is applied to the InceptionV3 and NASNetLarge models, which are processed with pre-trained weights to classify the input images. Then, blocks are cut from the points indicated in Fig. 4 in InceptionV3 and NASNetLarge models, and from the classification layer in the FINet model. As a result of block cutting, outputs with $12 \times 12 \times 768$ shapes in the InceptionV3 model, $14 \times 14 \times 2016$ shapes in the NASNetLarge model, and $7 \times 7 \times 192$ shapes in the FINet model are obtained. By adding new layers with different values consisting of convolution, batch normalization, ReLU activation function, average pooling, and dropout layers to all three models, a common output of $4 \times 4 \times 192$ shapes with the
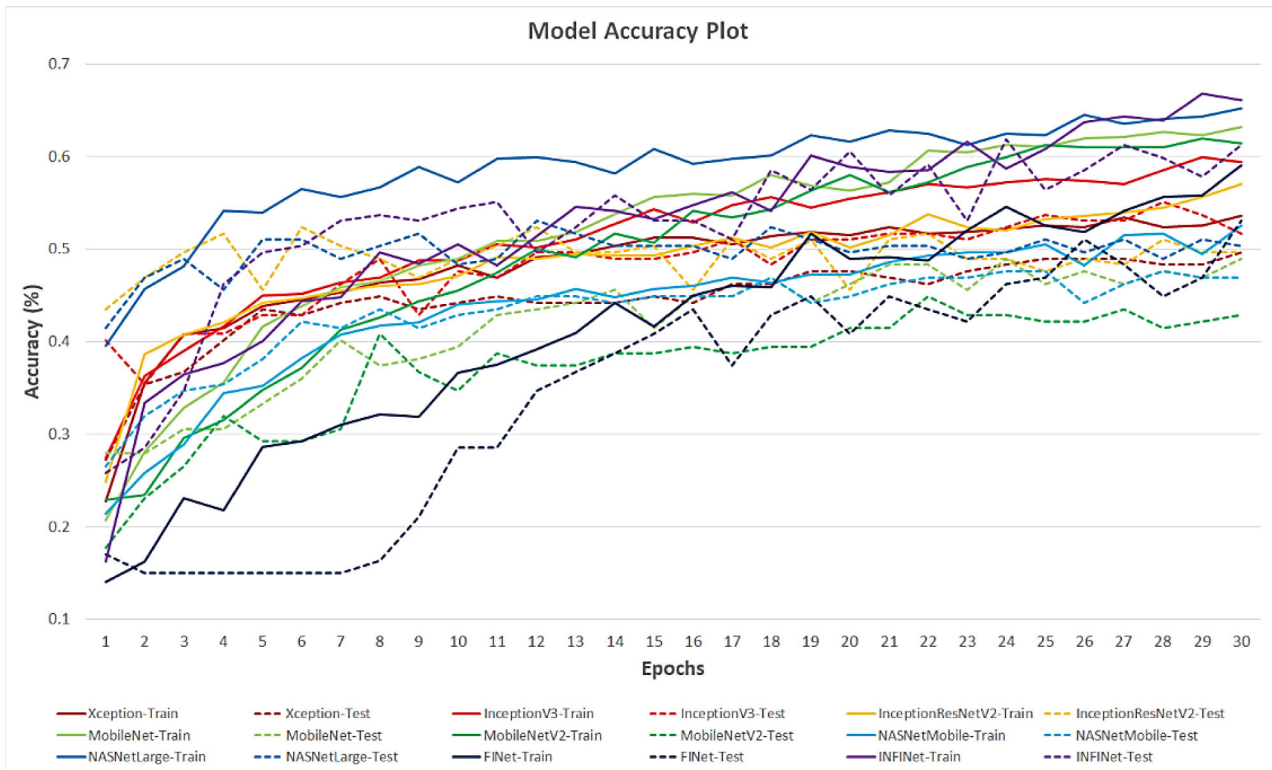
(a)
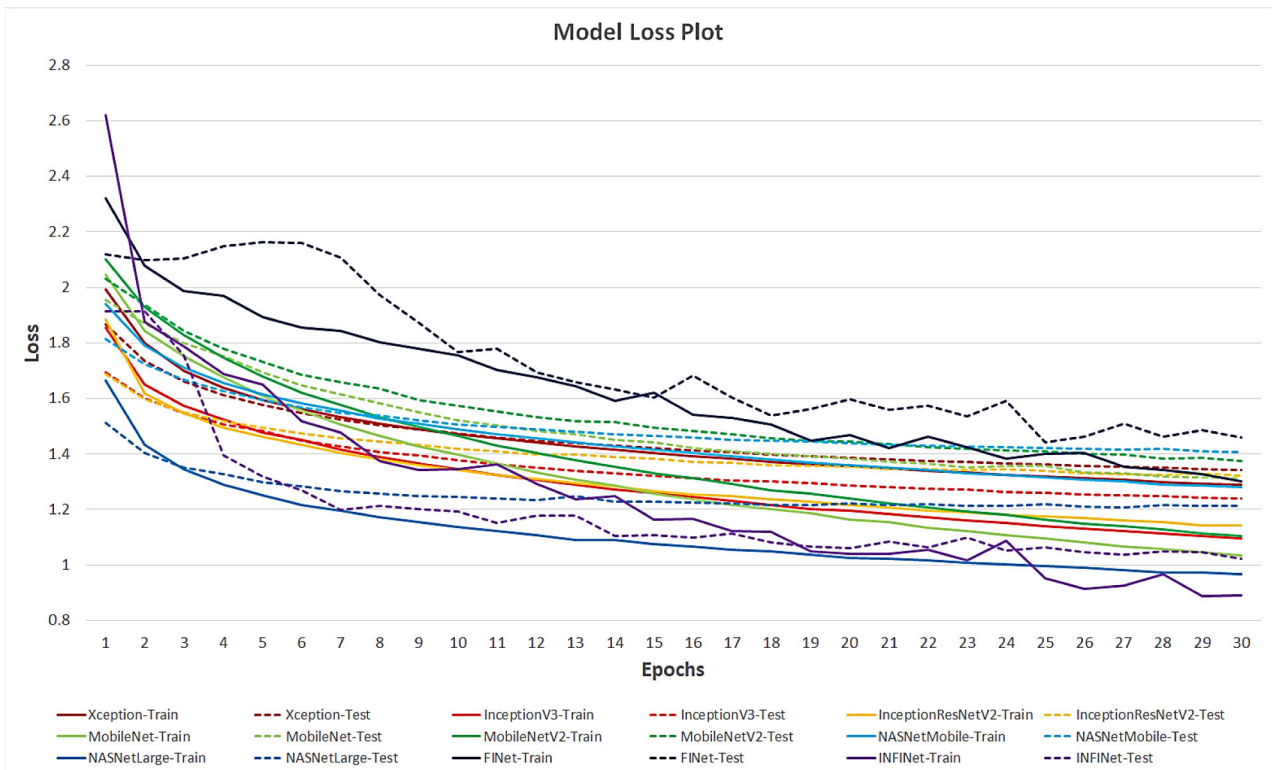
(b)

Fig. 6 Training plots of the models for age estimation in the FG-NET dataset: **a** accuracy and and **b** loss

Table 4 Numerical results obtained from model training for age estimation in the FG-NET dataset

Model	Accuracy	Loss	Precision	Recall	F1-Score
Xception	0.4966	1.3421	0.5350	0.4985	0.4469
InceptionV3	0.5170	1.2407	0.5131	0.5171	0.4904
InceptionV3_CutLayer	0.5442	1.1323	0.5066	0.5431	0.5190
InceptionResNetV2	0.4966	1.3207	0.4441	0.4963	0.4510
MobileNet	0.4896	1.3113	0.4726	0.4918	0.4776
MobileNetV2	0.4286	1.3731	0.3830	0.4229	0.3963
NASNetMobile	0.4694	1.4073	0.4144	0.4685	0.4225
NASNetLarge	0.5034	1.2118	0.4831	0.5061	0.4574
NASNetLarge_CutLayer	0.5238	1.1457	0.5099	0.5247	0.4864
FINet	0.5306	1.4579	0.5214	0.5286	0.5178
INFINet	0.6122	1.0234	0.6014	0.6129	0.6027

same feature values is obtained. After all three models have values of $4 \times 4 \times 192$, the models are concatenated. After the models are concatenated, layers consisting of convolution, batch normalization, ReLU activation function, average pooling, global average pooling, dropout, and softmax blocks are added to the end of the model, respectively. In this way, a new model with a structure consisting of these three models, which was not tried to be concatenated in another study before, has been created.

Each layer used in the INFINet model can be summarized as follows (the layers defined in the FINet model are not given again, only the layers that are different from the FINet model are defined):

- AvgPool: It was used to average the values in the area covered by the filter to reduce the number of parameters of the neural network by reducing the size of the feature map created by the convolution layer.
- Global average pool: It was used to reduce all spatial dimensions by averaging each feature map.

2.5 Evaluation metrics

After the model is created, various evaluation metrics are needed to measure how its performance works [55]. Evaluation metrics mostly come from the confusion matrix [56]. The four most common evaluation metrics (accuracy, precision, recall, and f1 score) were used for age and gender estimation on an imbalanced dataset of human face images. Mathematical expressions given in Eqs. 1, 2, 3, and 4 have been used to determine the accuracy, precision, recall, and f1 score of each model discussed in this study.

$$Precision = \frac{True\ Positive}{True\ Positive\ +\ False\ Positive} \quad (2)$$

$$Recall = \frac{True\ Positive}{True\ Positive\ +\ False\ Negative} \quad (3)$$

$$F1 - Score = 2 \times \frac{Precision\ \times\ Recall}{Precision\ +\ Recall} \quad (4)$$

The overall performance of the proposed models on the datasets was evaluated with the accuracy metric. In addition, the precision metric, used in cases where false positives are costly, was used to ensure that the proposed models correctly recognized a particular class, and the recall metric, used in cases where false negatives were costly, was used to measure the ability of the proposed models not to miss a class. The F1-score metric was used to measure the overall performance of the proposed models in a balanced way.

3 Results and discussion

In this study, training processes have been carried out with both the newly developed FINet model and the INFINet model, which was created as a result of the concatenation of the InceptionV3, NASNetLarge, and FINet models, for age and gender estimation using the imbalanced dataset. To compare the results obtained with state-of-the-art DCNN models in the field of AI, the same training was conducted with the Xception, InceptionV3, InceptionResNetV2, MobileNet, MobileNetV2, NASNetMobile, and NASNetLarge models. The training of the models was carried out
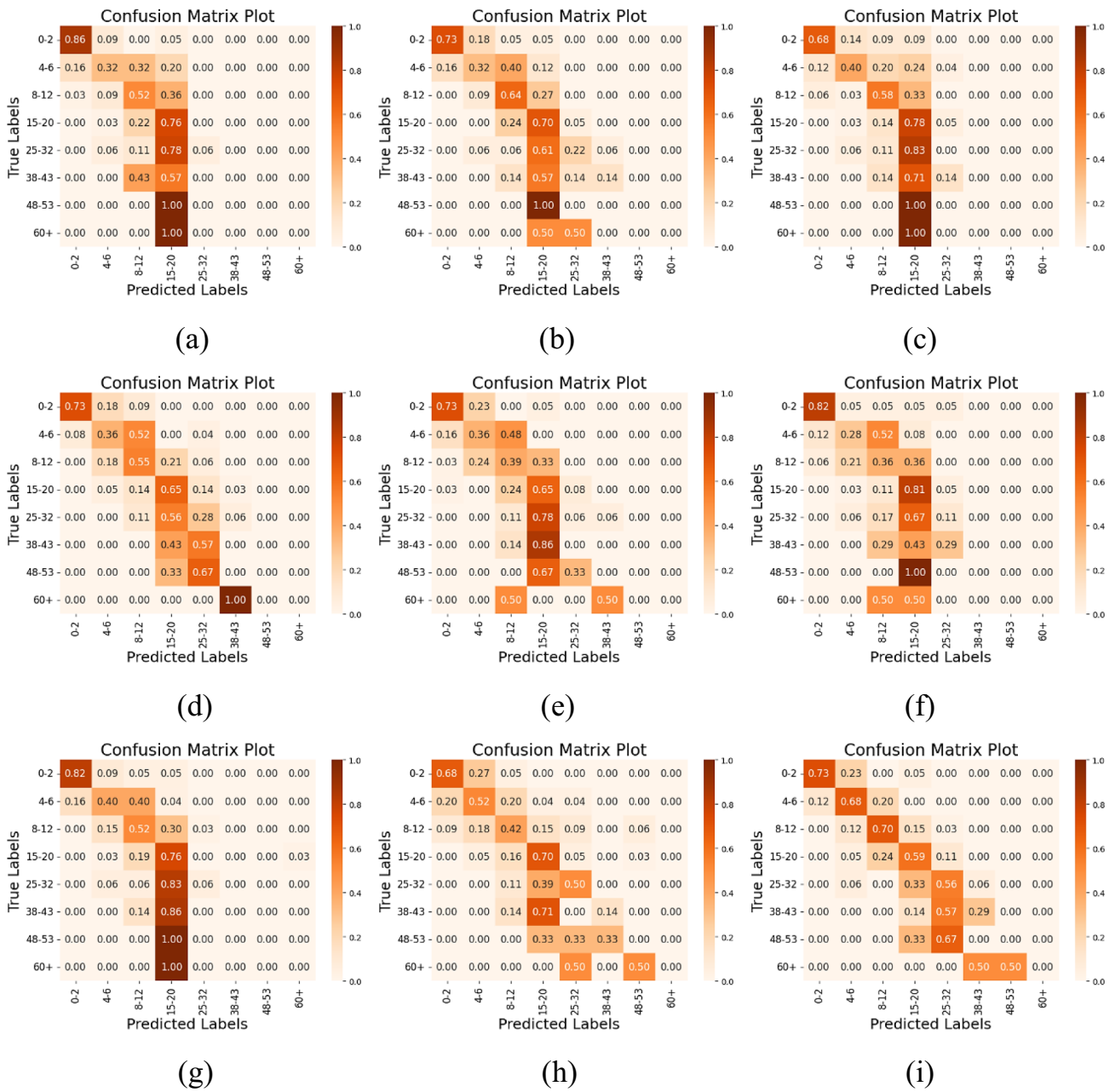
$$Accuracy = \frac{True\ Positive\ +\ True\ Negative}{True\ Positive\ +\ False\ Positive\ +\ False\ Negative\ +\ True\ Negative} \quad (1)$$

Fig. 7 Confusion matrix plots obtained from model training for age estimation in the FG-NET dataset: **a** Xception, **b** InceptionV3, **c** InceptionResNetV2, **d** MobileNet, **e** MobileNetV2, **f** NASNetMobile, **g** NASNetLarge, **h** FINet, and **i** INFINet

separately according to the train-test dataset, which was first separated according to the eight-class age groups and then according to the two-class gender groups.

## 3.1 Results for age estimation

The accuracy and loss graphs obtained because of the model training for age estimation are given in Fig. 6. In addition, the numerical results obtained from the models

for age estimation are given in Table 4, and the confusion matrix plots are given in Fig. 7.

When the accuracy and loss graphs given in Fig. 6, the numerical results given in Table 4, and the confusion matrix results given in Fig. 7 are examined, it is understood that the age estimation has been performed most successfully by the INFINet model. In the INFINet model, 61.22% accuracy and 1.0234 loss values were obtained. In Table 4, the accuracy rate of the InceptionV3 model, which was
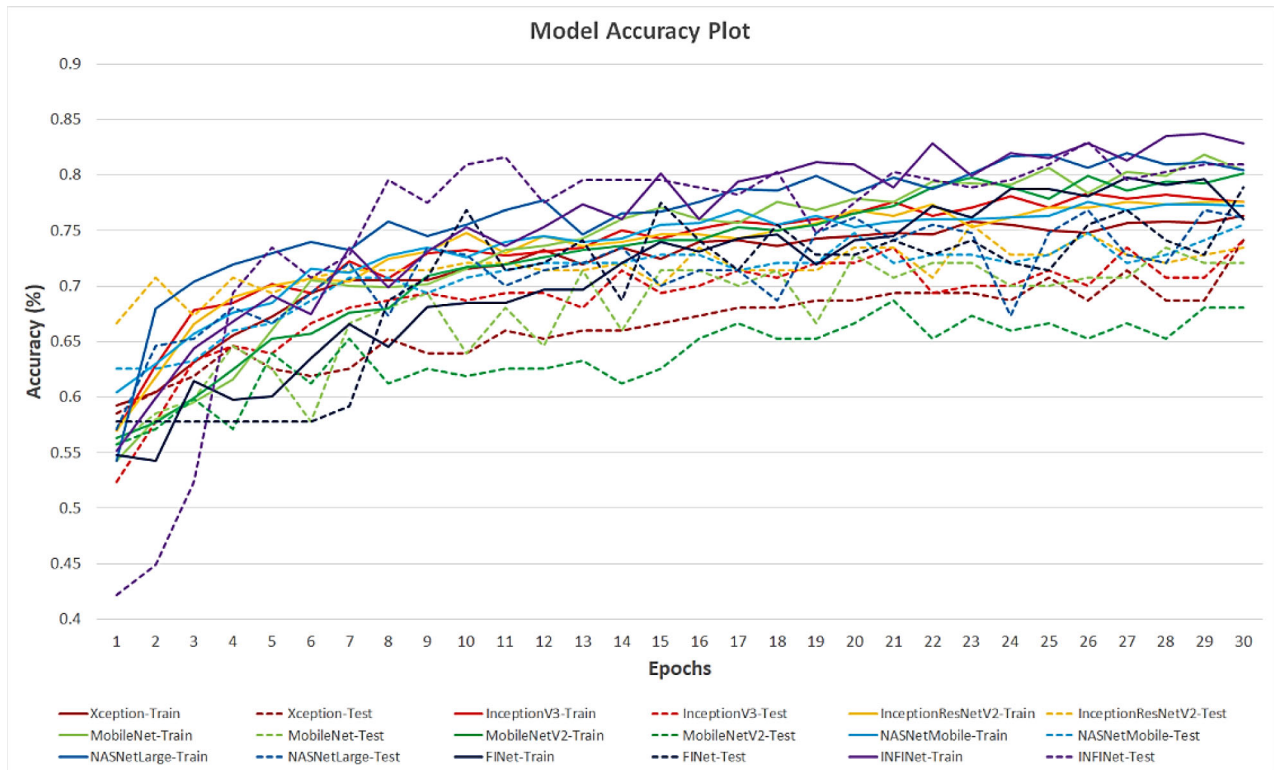
(a)
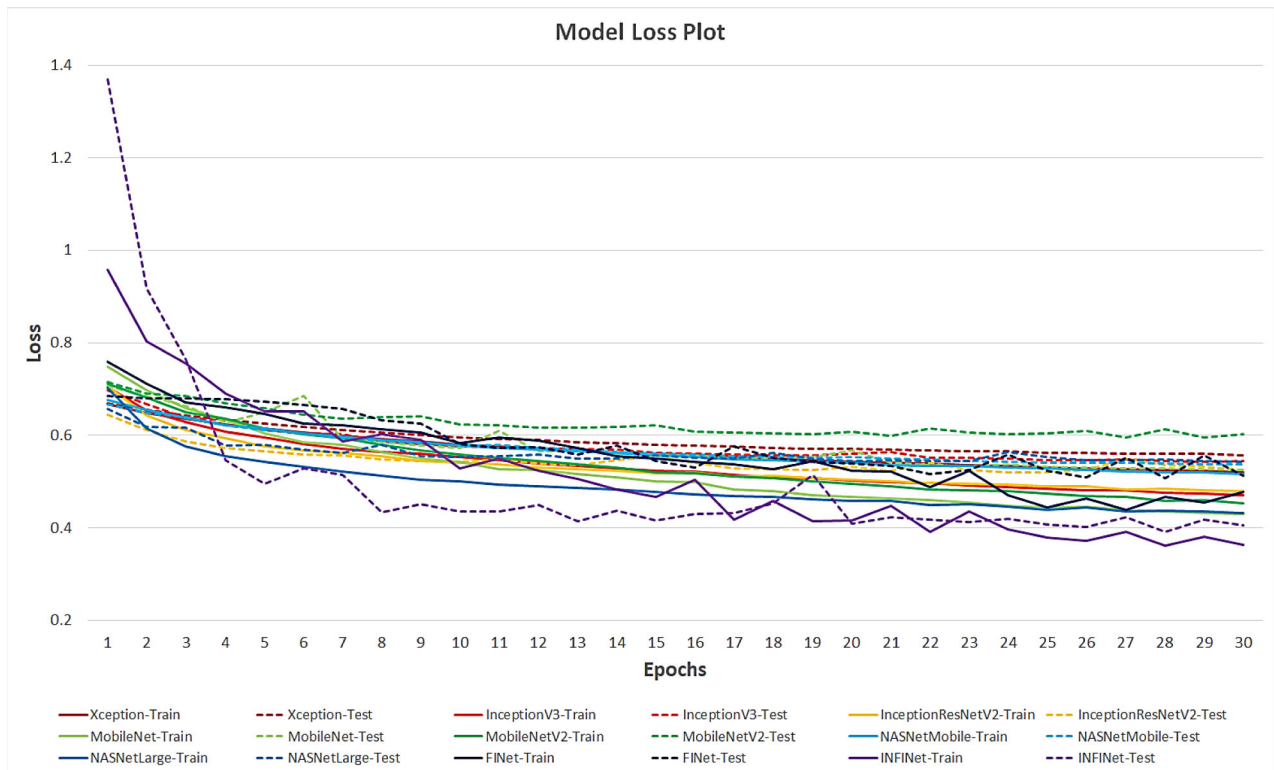
(b)

Fig. 8 Training plots of the models for gender estimation in the FG-NET dataset: **a** accuracy and **b** loss

Table 5 Numerical results obtained from model training for gender estimation in the FG-NET dataset

Model	Accuracy	Loss	Precision	Recall	F1-Score
Xception	0.7415	0.5575	0.7427	0.7393	0.7367
InceptionV3	0.7415	0.5446	0.7442	0.7419	0.7310
InceptionV3_CutLayer	0.7687	0.4925	0.7700	0.7714	0.7636
InceptionResNetV2	0.7347	0.5257	0.7411	0.7324	0.7225
MobileNet	0.7211	0.5252	0.7220	0.7205	0.7220
MobileNetV2	0.6803	0.6032	0.6910	0.6784	0.6847
NASNetMobile	0.7551	0.5382	0.7653	0.7508	0.7467
NASNetLarge	0.7619	0.5423	0.7605	0.7620	0.7663
NASNetLarge_CutLayer	0.7755	0.4684	0.7922	0.7739	0.7667
FINet	0.7891	0.5127	0.7969	0.7888	0.7836
INFINet	0.8095	0.4050	0.8105	0.8105	0.8105

51.70%, increased to 54.42% with the InceptionV3_CutLayer model obtained after block cutting. In addition, the accuracy rate of the NASNetLarge model, which was 50.34%, increased to 52.38% with the NASNetLarge_CutLayer model obtained after block cutting. Therefore, the block cutting process applied to the InceptionV3 and NASNetLarge models made a significant contribution to the successful results obtained with the INFINet model in age estimation. In addition, the FINet model with a success accuracy of 53.06% has been the best model after the INFINet model. When examined in terms of age groups, both INFINet and FINet models made the most inaccurate classification in the 48–53 and 60 + age groups. Both models performed better than other models in terms of age groups.

3.2 Results for gender estimation

The accuracy and loss graphs obtained because of the model training for gender estimation are given in Fig. 8. In addition, the numerical results obtained from the models for gender estimation are given in Table 5, and the confusion matrix plots are given in Fig. 9.

When the results given in Figs. 8, 9 and Table 5 are examined, it is seen that INFINet has been the most successful model in gender estimation. With the INFINet model, 80.95% accuracy and 0.4050 loss values were achieved. In Table 5, the accuracy rate of the InceptionV3 model, which was 74.15%, increased to 76.87% with the InceptionV3_CutLayer model obtained after block cutting. In addition, the accuracy rate of the NASNetLarge model, which was 76.19%, increased to 77.55% with the NASNetLarge_CutLayer model obtained after block cutting. Therefore, the block cutting process applied to the InceptionV3 and NASNetLarge models made a significant contribution to the successful results obtained with the

INFINet model in gender estimation. In addition, the FINet model with a success accuracy of 78.91% has been the best model after the INFINet model. As in age estimation, the two best models in gender estimation were seen as INFINet and FINet.

3.3 Comparative analysis

To better compare and discuss all the models discussed in this study, the datasets created for both age and gender estimation were trained separately under the same conditions. Comparative accuracy and loss graphs, other numerical results, and confusion matrix results obtained because of the training show that the proposed models in both age and gender estimation are better or close to state-of-the-art DCNN models in the field of AI. In addition to these remarkable successful results in the proposed models, the models should also be compared in terms of the number of parameters. In this direction, the total parameter numbers of the models discussed in this study are given in Fig. 10.

According to the total number of parameters given in Fig. 10, it is seen that the newly developed FINet model has the least number of parameters. The second-best accuracy in both age and gender estimation was achieved with the FINet model, which has the least complexity. For this reason, it is predicted that the FINet model can be competitive with state-of-the-art models in AI in age and gender estimation. Although the number of parameters of the INFINet model, which is created because of the concatenation of InceptionV3, NASNetLarge, and FINet models, is 12.06M more than the InceptionV3 model, it is 51.07M less than the NASNetLarge model. The most successful age and gender estimation was achieved with the INFINet model. Although it has 33.88M parameters, which is higher than FINet with 1.80M, MobileNetV2 with
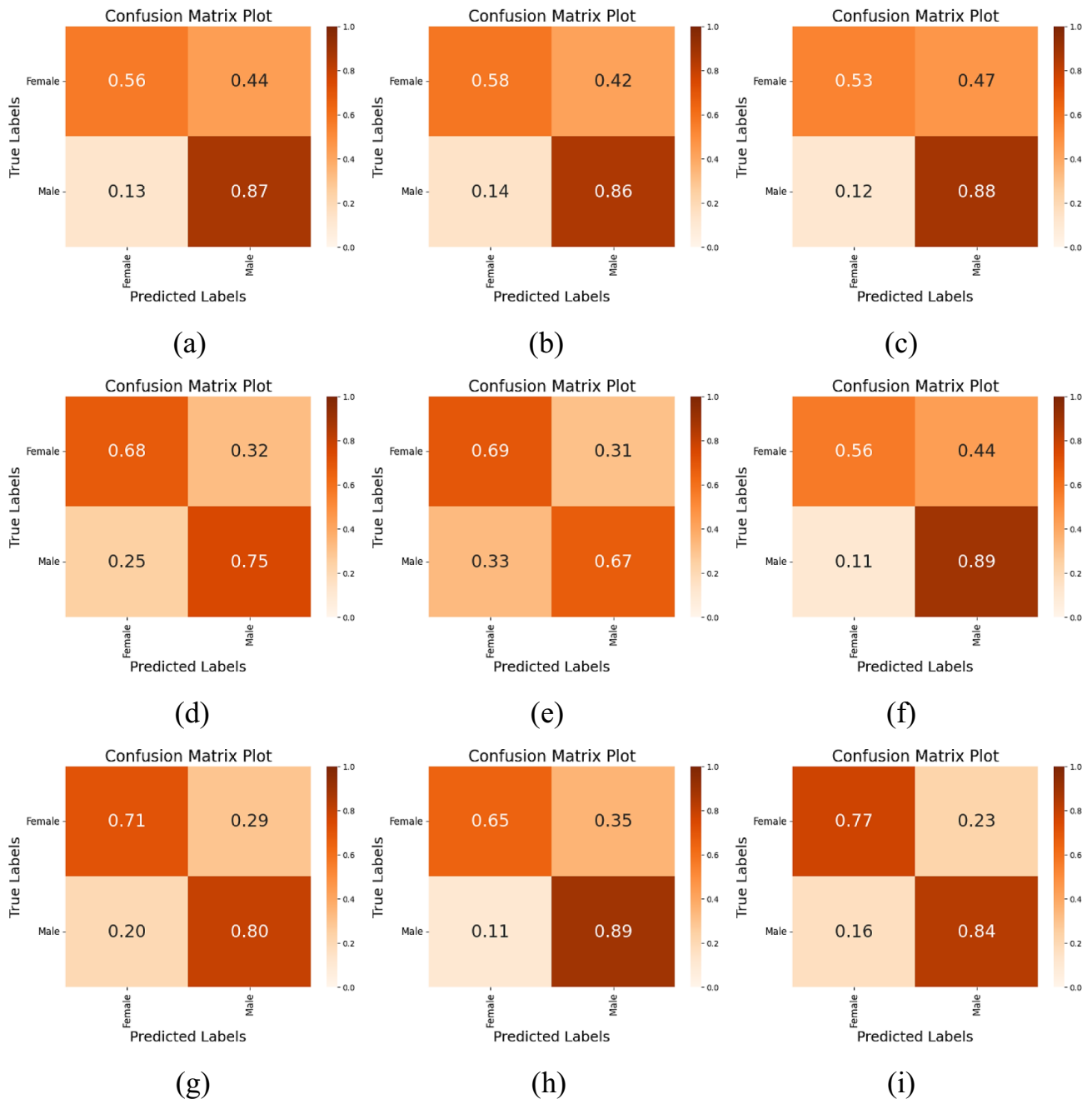
Fig. 9 Confusion matrix plots obtained from model training for gender estimation in the FG-NET dataset: **a** Xception, **b** InceptionV3, **c** InceptionResNetV2, **d** MobileNet, **e** MobileNetV2, **f** NASNetMobile, **g** NASNetLarge, **h** FINet, and **i** INFINet

2.27M, MobileNet with 3.24, NASNetMobile with 4.28M, Xception with 20.88M, and InceptionV3 with 21.82M, comparing the performance of the INFINet model with these models still makes the trade-off reasonable. To logically support such a claim, smaller models only achieved an overall accuracy of < 54% for age estimation and only < 79% for gender estimation, whereas the proposed INFINet model reached 61.22% for age estimation and

80.95% for gender estimation. In addition, compared to larger models such as InceptionResNetV2 (age: 49.66%, gender: 73.43%) and NASNetLarge (age: 50.34%, gender: 76.19%), the proposed INFINet model exceeds the performance reached by these models.

It is quite difficult to make high-success age and gender estimations in imbalanced datasets. Although the success rates are not very high, it is seen that the developed FINet
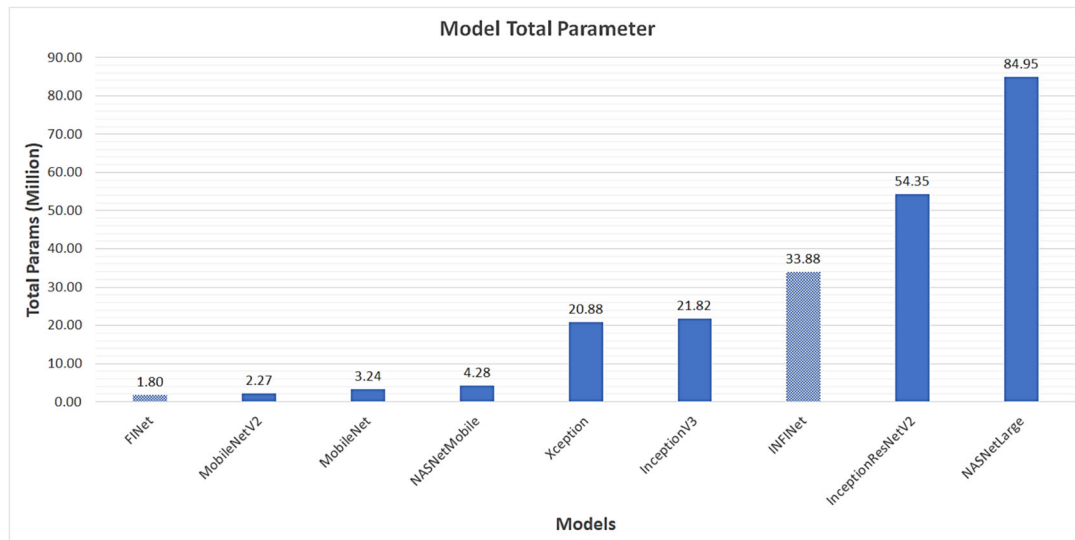
Fig. 10 Total number of parameters of the models

Table 6 Number of images of the train-test dataset obtained for each age and gender group in the UTKFace dataset

	Groups	Train (%80)	Test (%20)	Total (%100)
Age	0–2	1284	321	1605
	4–6	480	120	600
	8–12	624	156	780
	15–20	981	245	1226
	25–32	5417	1355	6772
	38–43	1338	334	1672
	48–53	1034	259	1293
	60 +	2152	538	2690
	Total	13,310	3328	16,638
Gender	Female	6472	1618	8090
	Male	6838	1710	8548
	Total	13,310	3328	16,638

and concatenated INFINet models are more successful than other models when considering the number of parameters in both age and gender estimation. This shows that the developed models are better than the compared models.

3.4 Testing models using different datasets

To test the models discussed in this study on different data, model training was carried out using a dataset called UTKFace. UTKFace is a large-scale dataset containing more than 20k + cropped face images with age, gender, and ethnicity. This dataset can be used in various tasks

such as age and gender estimation, face detection, and age progression/regression [61, 62]. Age groups are divided in this dataset as described in the previous section. Age groups are divided into eight classes 0–2, 4–6, 8–12, 15–20, 25–32, 38–53, 48–53, and 60 + , and gender groups are divided into two classes female and male. In Table 6, the number of images in each class according to age and gender groups is given.

After applying the same pre-processing described in the previous sections of the study to this dataset, all model training was carried out. Numerical results obtained from the models as a result of extensive experiments using the UTKFace dataset are given in detail in Table 7. In addition, confusion matrix plots obtained from models for age and gender estimation are given in Figs. 11 and 12, respectively.

When the numerical results given in Table 7, Figs. 11 and 12 are examined, the proposed INFINet model was the most successful compared to other models in both age and gender estimation, even when a different dataset was used. The INFINet model achieved 72.00% accuracy, 0.7408 loss in age estimation, and 90.50% accuracy, 0.2274 loss in gender estimation on the UTKFace dataset. It is seen that the proposed INFINet model is successful not only on a single dataset but also on a different datasets. In terms of age groups, the INFINet model made the most inaccurate classification in the 38–43 and 48–53 age groups. The proposed FINet model achieved the best success rate after the INFINet model in both age and gender estimation. This shows that the FINet model is competitive with other models in age and gender estimation. Therefore, the pro-

Table 7 Numerical results obtained from models using the UTKFace dataset

	Model	Accuracy	Loss	Precision	Recall	F1-Score
Age	Xception	0.6439	1.0291	0.5937	0.6431	0.5740
	InceptionV3	0.6391	1.0062	0.5913	0.6411	0.5924
	InceptionResNetV2	0.6409	0.9918	0.5852	0.6400	0.5820
	MobileNet	0.6535	0.9589	0.6019	0.6525	0.6049
	MobileNetV2	0.6478	1.0210	0.6018	0.6488	0.6004
	NASNetMobile	0.6154	1.0883	0.5385	0.6161	0.5461
	NASNetLarge	0.6340	1.0157	0.5731	0.6360	0.5802
	FINet	0.6986	0.8502	0.6638	0.6984	0.6560
	INFINet	0.7200	0.7408	0.6877	0.7208	0.6885
Gender	Xception	0.8368	0.3595	0.8397	0.8357	0.8351
	InceptionV3	0.8380	0.3511	0.8349	0.8354	0.8351
	InceptionResNetV2	0.8453	0.3494	0.8457	0.8443	0.8449
	MobileNet	0.8540	0.3279	0.8549	0.8554	0.8551
	MobileNetV2	0.8215	0.3919	0.8243	0.8217	0.8203
	NASNetMobile	0.8248	0.3708	0.8249	0.8254	0.8251
	NASNetLarge	0.8585	0.3239	0.8597	0.8606	0.8551
	FINet	0.8810	0.2655	0.8846	0.8808	0.8803
	INFINet	0.9050	0.2274	0.9092	0.9062	0.9051

posed FINet and INFINet models were introduced to the literature as new models.

3.5 Discussion

In the studies conducted with the FG-NET dataset in the literature review in the introduction chapter of this study, the dataset was used in training and classified according to different age groups such as 2–4, and successful results were achieved [34–37]. High successes have also been achieved in studies in the literature where training is made with different datasets and only testing is done with the FG-NET dataset [38–40]. Additionally, there are other studies using datasets other than the FG-NET dataset [28–30, 33, 46]. In our study, the FG-NET dataset was divided into eight groups. Moreover, it was used not only in the testing process but also in the training–testing processes. In this regard, recent works in which training–testing processes were performed with the FG-NET dataset and age groups were used similar to our work were examined, and the results are presented comparatively in Table 8.

When Table 8 is examined, it can be seen that the proposed INFINet model performs much better than other methods. We can say that the proposed FINet model exhibits a competitive performance with other methods.

4 Conclusions

In this study, an imbalanced dataset of human face images was used for age and gender estimation. Two new DCNN models have been developed for age and gender estimation with an imbalanced dataset of human face images. The first of the developed models, the FINet model, is a new model design with a parametrically modifiable structure. The second developed model, the INFINet model, is a model created because of the concatenation of InceptionV3, NASNetLarge, and FINet models after improvements. Both models were designed for the first time and have unique structures.

The FINet and INFINet models developed for age and gender estimation have been compared with many models that have shown significant success in the field of AI in recent years. FINet, INFINet, Xception, InceptionV3, InceptionResNetV2, MobileNet, MobileNetV2, NASNet-Mobile, and NASNetLarge models were trained under the same conditions and compared. As a result of the comparisons, the highest accuracy (age: %61.22, gender: %80.95 in the FG-NET dataset, age: 72.00%, gender: 90.50% in the UTKFace dataset) and lowest loss (age: 1.0234, gender: 0.4050 in the FG-NET dataset, age: 0.7408, gender: 0.2274 in the UTKFace dataset) values were achieved with the INFINet model developed for both age and gender estimation. It is one of the important achievements of the study that the INFINet model, which is
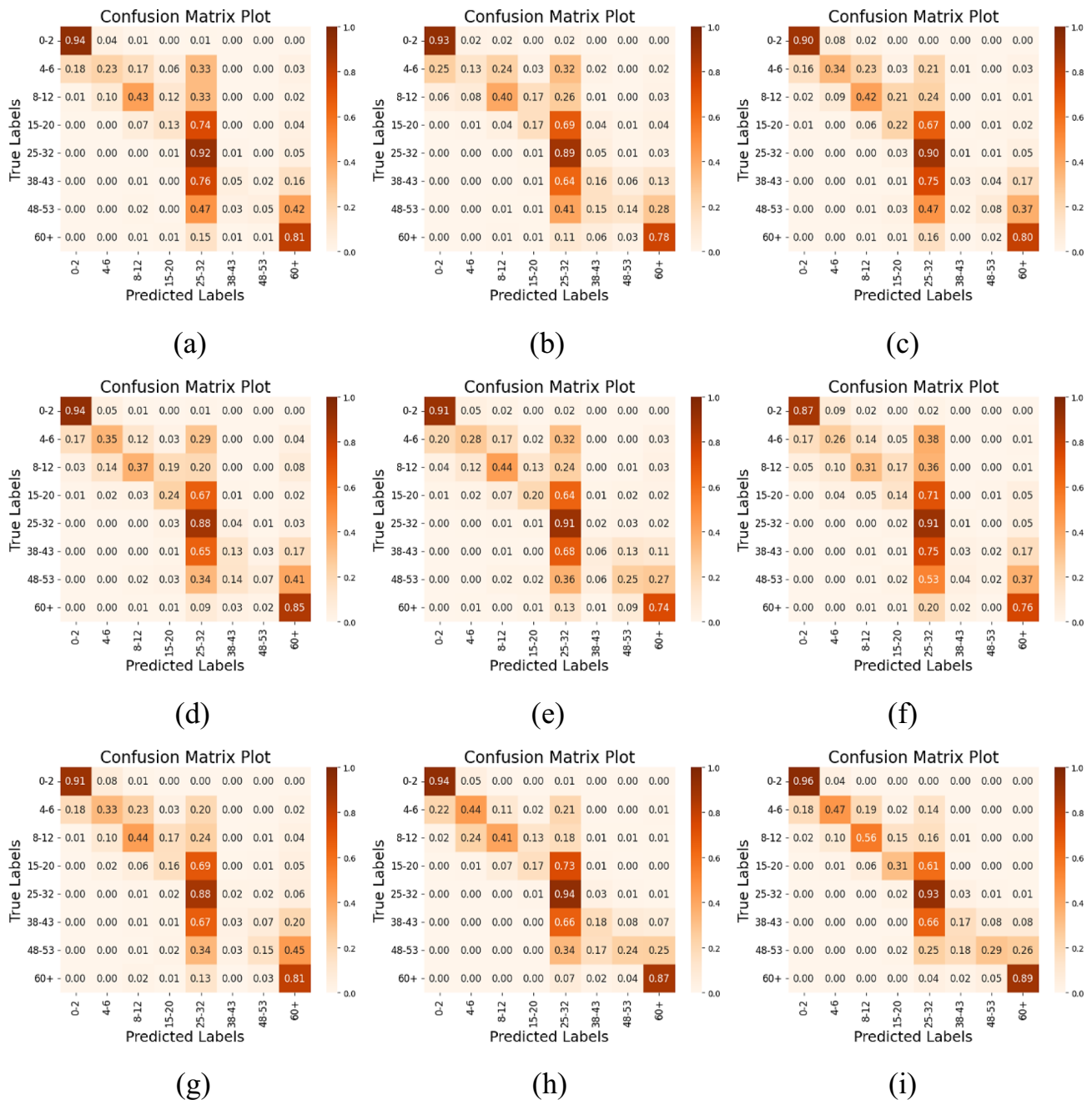
Fig. 11 Confusion matrix plots obtained from models for age estimation using the UTKFace dataset: **a** Xception, **b** InceptionV3, **c** InceptionResNetV2, **d** MobileNet, **e** MobileNetV2, **f** NASNetMobile, **g** NASNetLarge, **h** FINet, and **i** INFINet

brought to a parameter between the combined InceptionV3 model and the NASNetLarge model, has achieved higher success than both these two models and other AI technologies. Another remarkable achievement is that with the FINet model, which has much fewer parameters than all the AI technologies discussed in this study, more successful accuracy values are achieved than all other models except for INFINet.

It has been concluded that FINet and INFINet models developed for age and gender estimation are competitive with other models in cases where it is difficult to obtain successful results with imbalanced datasets. Future studies planned to be carried out include several objectives:

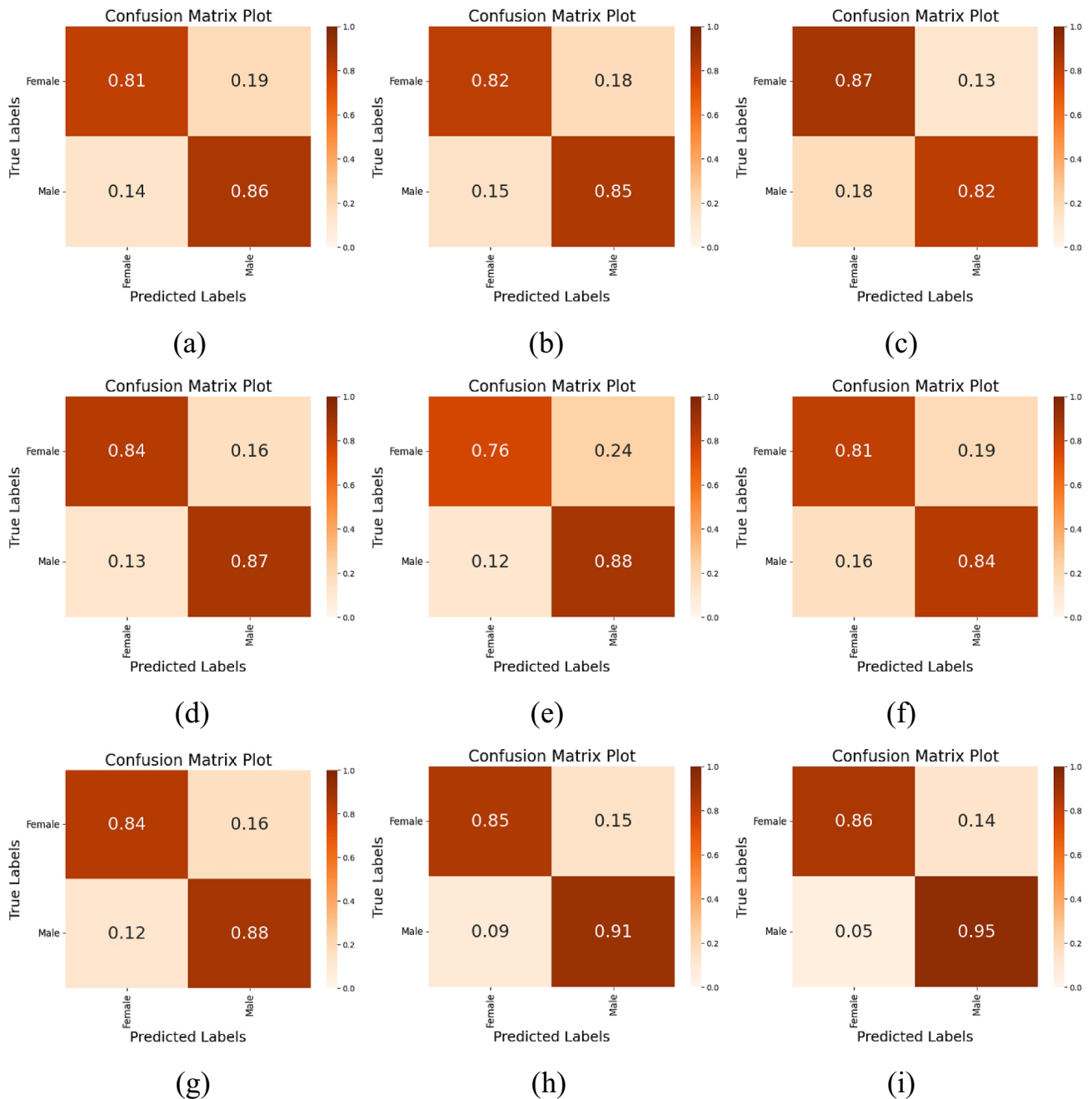- Application of the developed FINet model with different parameters

Fig. 12 Confusion matrix plots obtained from models for gender estimation using the UTKFace dataset: **a** Xception, **b** InceptionV3, **c** InceptionResNetV2, **d** MobileNet, **e** MobileNetV2, **f** NASNetMobile, **g** NASNetLarge, **h** FINet, and **i** INFINet

- Developing the FINet model and producing its second version
- Creating new and better models by combining the FINet model with different AI technologies
- Estimating age on the selected gender dataset, after determining the gender
- Estimating age based on specific ages
- Creating models for object recognition using the Transformer approach

- Testing the developed models on many different balanced/imbalanced datasets
- Evaluating model performance by applying balancing techniques and augmentation methods
- Application of the developed models to real-world problems

Table 8 Comparison of the proposed work with recent works

Dataset	References	Method	Accuracy	
			Age	Gender
FG-NET	Dammak et al. [57]	Deep learning based method	54.66%	–
	Abbes et al. [58]	Landmarks-based features	56.46%	–
	Proposed work	FINet	53.06%	78.91%
	Proposed work	INFINet	61.22%	80.95%
UTKFace	Jamoliddin and Yoo [59]	CNN-based classification	61.34%	88.67%
	Apuandi et al. [60]	ConvELM	71.17%	–
	Proposed work	FINet	69.86%	88.10%
	Proposed work	INFINet	72.00%	90.50%

Declarations

Conflict of interest No conflict of interest or common interest has been declared by the author.

Ethical approval This study does not require ethics committee permission or any special permission.

References

1. Gupta SK, Nain N (2023) Single attribute and multi attribute facial gender and age estimation. Multimed Tools Appl 82(1):1289–1311. https://doi.org/10.1007/s11042-022-12678-6
2. Sharma N, Sharma R, Jindal N (2022) Face-based age and gender estimation using improved convolutional neural network approach. Wireless Pers Commun 124(4):3035–3054. https://doi.org/10.1007/s11277-022-09501-8
3. Saraswat M, Gupta P, Yadav RP, Yadav R and Sonkar S (2022, July) Age, gender and emotion estimation using deep learning. In: Congress on intelligent systems: Proceedings of CIS 2021, Singapore: Springer Nature Singapore, vol 2, pp 59–70. https://doi.org/10.1007/978-981-16-9113-3_6
4. Dammak S, Mliki H, Fendri E (2023) Gender estimation based on deep learned and handcrafted features in an uncontrolled environment. Multimedia Syst 29(1):421–433. https://doi.org/10.1007/s00530-022-01011-8
5. Zhang B, Bao Y (2022) Cross-dataset learning for age estimation. IEEE Access 10:24048–24055. https://doi.org/10.1109/ACCESS.2022.3154403
6. Vidyarthi P, Dhavale S and Kumar S (2022, August) Gender and age estimation using transfer learning with multi-tasking approach. In: 2022 2nd Asian conference on innovation in technology (ASIANCON), IEEE, pp 1–5. https://doi.org/10.1109/ASIANCON55314.2022.9908952
7. Nawaz Y, Arif MS, Shatanawi W, Nazeer A (2021) An explicit fourth-order compact numerical scheme for heat transfer of boundary layer flow. Energies 14(12):3396. https://doi.org/10.3390/en14123396
8. Nawaz Y, Arif MS, Abodayeh K (2022) A third-order two-stage numerical scheme for fractional Stokes problems: a comparative computational study. J Comput Nonlinear Dyn 7(10):101004. https://doi.org/10.1115/1.4054800
9. Nawaz Y, Arif MS, Abodayeh K (2022) An explicit-implicit numerical scheme for time fractional boundary layer flows. Int J Numer Meth Fluids 94(7):920–940. https://doi.org/10.1002/fld.5078
10. Akgül İ (2023) Mobile-DenseNet: detection of building concrete surface cracks using a new fusion technique based on deep learning. Heliyon 9(10):e21097. https://doi.org/10.1016/j.heliyon.2023.e21097
11. Wu Y, Zhao S, Xing Z, Wei Z, Li Y, Li Y (2023) Detection of foreign objects intrusion into transmission lines using diverse generation model. IEEE Trans Power Deliv. https://doi.org/10.1109/TPWRD.2023.3279891
12. Xing Z, Zhao S, Guo W, Meng F, Guo X, Wang S, He H (2023) Coal resources under carbon peak: Segmentation of massive laser point clouds for coal mining in underground dusty environments using integrated graph deep learning model. Energy 285:128771. https://doi.org/10.1016/j.energy.2023.128771
13. Arif MS, Mukheimer A, Asif D (2023) Enhancing the early detection of chronic kidney disease: a robust machine learning model. Big Data Cogn Comput 7(3):144. https://doi.org/10.3390/bdcc7030144
14. Asif D, Bibi M, Arif MS, Mukheimer A (2023) Enhancing heart disease prediction through ensemble learning techniques with hyperparameter optimization. Algorithms 16(6):308. https://doi.org/10.3390/a16060308
15. Thaneeshan R, Thanikasalam K and Pinidiyaarachchi A (2022, December) Gender and age estimation from facial images using deep learning. In: 2022 7th International conference on information technology research (ICITR), IEEE, pp 1–6. https://doi.org/10.1109/ICITR57877.2022.9993277
16. Shi C, Zhao S, Zhang K, Feng X (2023) Multi-task multi-scale attention learning-based facial age estimation. IET Signal Process 17(2):e12190. https://doi.org/10.1049/sil2.12190
17. Di Mascio T, Fantozzi P, Laura L and Rughetti V (2022) Age and gender (face) recognition: a brief survey. In: Methodologies and intelligent systems for technology enhanced learning, 11th international conference, Springer International Publishing, vol 11, pp 105–113. https://doi.org/10.1007/978-3-030-86618-1_11
18. Kulkarni MA, Joshi MP, Sindgi MS, Rakshasbhuvankar MS, Kumar MV, Dachawar M (2022) Detection of gender and age using machine learning. Int J Res Appl Sci Eng Technol 10:1537–1542. https://doi.org/10.22214/ijraset.2022.48268

19. Arya S, Khan M, Agarwal A, Gaur A, Mallick B (2022) Age estimation and gender recognition technique using deep learning. Int J Res Appl Sci Eng Technol 10:326–331. https://doi.org/10.22214/ijraset.2022.42145

20. Tariq MU, Akram A, Yaqoob S, Rasheed M and Ali MS (2022) Real time age and gender classification using Vgg19. Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition), vol 41(12), pp 641–655. https://doi.org/10.17605/OSF.IO/BKJWH

21. Lee SH, Hosseini S, Kwon HJ, Moon J, Koo HI and Cho NI (2018, January) Age and gender estimation using deep residual learning network. In: 2018 International workshop on advanced image technology (IWAIT), IEEE, pp 1–3. https://doi.org/10.1109/IWAIT.2018.8369763

22. Terhörst P, Huber M, Kolf JN, Zelch I, Damer N, Kirchbuchner F and Kuijper A (2019, September) Reliable age and gender estimation from face images: stating the confidence of model predictions. In: 2019 IEEE 10th international conference on biometrics theory, applications and systems (BTAS), IEEE, pp 1–8. https://doi.org/10.1109/BTAS46853.2019.9185975

23. Ozbulak G, Aytar Y and Ekenel HK (2016, September) How transferable are CNN-based features for age and gender classification?. In: 2016 International conference of the biometrics special interest group (BIOSIG), IEEE, pp 1–6. https://doi.org/10.1109/BIOSIG.2016.7736925

24. Mallouh AA, Qawaqneh Z, Barkana BD (2019) Utilizing CNNs and transfer learning of pre-trained models for age range classification from unconstrained face images. Image Vis Comput 88:41–51. https://doi.org/10.1016/j.imavis.2019.05.001

25. Duan M, Li K, Yang C, Li K (2018) A hybrid deep learning CNN–ELM for age and gender classification. Neurocomputing 275:448–461. https://doi.org/10.1016/j.neucom.2017.08.062

26. Chen BC, Chen CS and Hsu WH (2014) Cross-age reference coding for age-invariant face recognition and retrieval. In: Computer vision–ECCV 2014: 13th European conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part VI 13. Springer International Publishing, pp 768–783

27. Rwigema J, Mfitumukiza J, Tae-Yong K (2021) A hybrid approach of neural networks for age and gender classification through decision fusion. Biomed Signal Process Control 66:102459. https://doi.org/10.1016/j.bspc.2021.102459

28. Yudin D, Shchendrygin M, Dolzhenko A (2020) Age and gender recognition on imbalanced dataset of face images with deep learning. In: Kovalev S, Tarassov V, Snasel V, Sukhanov A (eds) Proceedings of the fourth international scientific conference "Intelligent Information Technologies for Industry" (IITI'19). IITI 2019. Advances in Intelligent Systems and Computing, vol 1156. Springer, Cham. https://doi.org/10.1007/978-3-030-50097-9_4

29. Sheoran V, Joshi S, Bhayani TR (2021) Age and gender prediction using deep CNNs and transfer learning. In: Singh SK, Roy P, Raman B, Nagabhushan P (eds) Computer vision and image processing. CVIP 2020. Communications in computer and information science, vol 1377. Springer, Singapore. https://doi.org/10.1007/978-981-16-1092-9_25

30. Alsaleh A, Perkgoz C (2023) A space and time efficient convolutional neural network for age group estimation from facial images. PeerJ Comput Sci 9:e1395. https://doi.org/10.7717/peerj-cs.1395

31. Medium (2024a) Gender and Age Detection using with Keras Tensorflow | Image Processing. Accessed in May 15, 2024 from https://medium.com/@ilaslanduzgun/gender-and-age-detection-using-with-keras-tensorflow-image-processing-90d7804473f8

32. Medium (2024b) NeuroNuggets: Age and Gender Estimation. Accessed in May 15, 2024 from https://medium.com/neuromation-blog/neuronuggets-age-and-gender-estimation-2807b1307a13

33. George G, Uppin C, Bello UA (2024) Human age estimation from face images with deep convolutional neural networks using transfer learning. Preprints 2024, 2024010350. https://doi.org/10.20944/preprints202401.0350.v1

34. Nithyashri J and Kulanthaivel G (2012, December) Classification of human age based on neural network using FG-NET aging database and wavelets. In: 2012 fourth international conference on advanced computing (ICoAC), IEEE, pp 1–5. https://doi.org/10.1109/ICoAC.2012.6416855

35. Choobeh AK (2013) An image-based method of distinguishing children from adults. Int J Inf Electron Eng 3(5):533. https://doi.org/10.7763/IJIEE.2013.V3.372

36. Samadi A and Pourghassem H (2013, April) Children detection algorithm based on statistical models and LDA in human face images. In: 2013 International Conference on Communication Systems and Network Technologies, IEEE, pp 206–209. https://doi.org/10.1109/CSNT.2013.52

37. Razalli H, Rahmat RWO, Khalid F, Sulaiman PS (2017) An image-based children age range verification and classification based on facial features angle distribution and face shape elliptical ratio. Adv Sci Lett 23(5):4026–4030. https://doi.org/10.1166/asl.2017.8271

38. Zhang Z, Song Y and Qi H (2017) Age progression/regression by conditional adversarial autoencoder. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 5810–5818. https://doi.org/10.1109/cvpr.2017.463

39. Tyagi M, Sood S (2017) Age group estimation using machine learning techniques: a review. Int J Eng Sci Res Technol 6(12):599–606. https://doi.org/10.5281/zenodo.1130907

40. Kumar S, Singh S, Kumar J, Prasad KMVV (2022) Age and gender classification using Seg-Net based architecture and machine learning. Multimed Tools Appl 81(29):42285–42308. https://doi.org/10.1007/s11042-021-11499-3

41. Cootes, T. (2014). FG-Net face and gesture recognition network

42. Github (2023a) Accessed in January 20, 2023 from https://yanweifu.github.io/FG_NET_data/

43. Kaggle (2023a) Accessed in January, 20, 2023 from https://www.kaggle.com/datasets/mulukentesfaye/fgnet

44. Garain A, Ray B, Singh PK, Ahmadian A, Senu N, Sarkar R (2021) GRA_Net: a deep learning model for classification of age and gender from facial images. IEEE Access 9:85672–85689. https://doi.org/10.1109/ACCESS.2021.3085971

45. Bukar AM, Ugail H, Connah D (2016) Automatic age and gender classification using supervised appearance model. J Electron Imaging 25(6):061605–061605

46. Levi G and Hassner T (2015) Age and gender classification using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 34–42)

47. Colab (2023) Google colaboratory. Retrieved in January, 25, 2023 from https://colab.research.google.com

48. Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1251–1258

49. Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826

50. Szegedy C, Ioffe S, Vanhoucke V and Alemi A (2017, February) Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI conference on artificial intelligence, vol 31, No. 1. https://doi.org/10.1609/aaai.v31i1.11231

51. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T and Adam H (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications. *arXiv preprint* arXiv:1704.04861

52. Sandler M, Howard A, Zhu M, Zhmoginov A and Chen LC (2018) Mobilenetv2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4510–4520

53. Zoph B, Vasudevan V, Shlens J and Le QV (2018) Learning transferable architectures for scalable image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 8697–8710

54. Montalbo FJP (2022) Diagnosing gastrointestinal diseases from endoscopy images through a multi-fused CNN with auxiliary layers, alpha dropouts, and a fusion residual block. Biomed Signal Process Control 76:103683. https://doi.org/10.1016/j.bspc.2022.103683

55. Sohan MF, Jabiullah MI, Rahman SSMM and Mahmud SH (2019, July) Assessing the effect of imbalanced learning on cross-project software defect prediction. In: 2019 10th International conference on computing, communication and networking technologies (ICCCNT), IEEE, pp 1–6. https://doi.org/10.1109/ICCCNT45670.2019.8944622

56. Zhang F, Zheng Q, Zou Y and Hassan AE (2016, May) Cross-project defect prediction using a connectivity-based unsupervised classifier. In: Proceedings of the 38th international conference on software engineering, pp 309–320. https://doi.org/10.1145/2884781.2884839

57. Dammak S, Mliki H, Fendri E (2021) Gender effect on age classification in an unconstrained environment. Multimed Tools App 80(18):28001–28014. https://doi.org/10.1007/s11042-021-11060-2

58. Abbes A, Ouarda W, Ayed YB (2024) Age-API: are landmarks-based features still distinctive for invariant facial age recognition? Multimed Tools Appl. https://doi.org/10.1007/s11042-024-18227-7

59. Jamoliddin U, Yoo JH (2022) Age and gender classification with small scale cnn. J Korea Inst Electron Commun Sci 17(1):99–104. https://doi.org/10.13067/JKIECS.2022.17.1.99

60. Apuandi I, Rachmawati E and Kosala G (2023, February) ConvELM: exploiting extreme learning machine on convolutional neural network for age estimation. In: 2023 International conference on artificial intelligence in information and communication (ICAIIC), IEEE, pp 407–412. https://doi.org/10.1109/ICAIIC57133.2023.10067115

61. Github (2023b) Accessed in September 21, 2023 from https://susanqq.github.io/UTKFace/

62. Kaggle (2023b) Accessed in September, 21, 2023 from https://www.kaggle.com/datasets/jangedoo/utkface-new