Age Group Classification using Convolutional Neural Network (CNN)

Muhammad Firdaus Mustapha¹, Nur Maisarah Mohamad ¹*, Ghazali Osman¹ and Siti Haslini Ab Hamid²

¹Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Kelantan, Bukit Ilmu, 18500 Machang, Kelantan, Malaysia
²FH Training Center, 16800 Pasir Puteh, Kelantan, Malaysia

*maisarah97.mohamad@gmail.com

Abstract. Age group classification is a complex task that is used to classify facial images or videos into predetermined age categories. It is an important task due to its numerous applications such as health, security, authentication system, recruitment, and also in intelligent social robots. Convolutional Neural Network (CNN) has recently shown excellent performance in analysing human face images and videos. This paper proposed an age group classification task using CNN that trained and tested with an All-Age Face (AAF) dataset. FaceNet deep learning model that uses CNN was applied in this study to compute a 128-d embedding that quantifies the face of the age group. The experiment included two age groups: Adolescence and Mature Adulthood. The proposed age group classification model achieved 84.90% accuracy for the training images and 85.12% accuracy for the test images. The experimental results showed that CNN is capable of achieving competitive classification accuracy throughout two age groups in the AAF dataset with unbalanced data distribution.

Keywords: age group classification, convolutional neural network, deep learning, still image face recognition

1. Introduction

Our identity is closely linked to our face. It is crucial in all activities involving human-to-human interaction. The automatic detection of the face and its characteristics have several potential uses. It may potentially become a hands-free option in the future or an amusing feature in creative games as an outcome of the digitalization connection between computer and human [1]. Security control, surveillance monitoring, and targeted marketing systems have all lately included automatic face detection, tracking, and classification into their systems [2]. Implementing an automatic face identification and classifier on a tracking platform can expand its use to a larger range of mobile services applications [3]. Several approaches and algorithms have been employed in the field of image processing and computer vision in recent years, employing popular approaches such as Principal Component Analysis (PCA) [4], Histogram Analysis [5], Convolutional Neural Networks (CNN) [6], and many more. Some of the techniques generate consistent findings for an investigated individual item attributes, while others recognize a specific object's appearance, such as colour orientation, object form, and etc. Face recognition algorithms have a variety of practical uses, including security platforms that control
the amount of people who may enter the airports, office buildings, and other high-security areas [7]. Besides, facial recognition technology is also used in autonomous systems for intelligent human-computer interactions, multimedia gaming and entertainment, accident prevention, and automatic age estimation [1].

The goal of age estimation is to apply a label to a facial image that indicates its exact age or the age group to which it belongs [8]. Most of the previous age estimation methods were designed to provide an accurate estimate of the real age. However, due to the specialization of aging effects on the face, determining an exact age is difficult. Different people age in diverse situations, which are impacted by a variety of internal and external factors such as ethnicity, heredity, health, environment, and lifestyle. Therefore, aging is uncontrolled, making age estimation difficult for humans in some instances and significantly more difficult for computer vision systems. Moreover, other aspects such as facial expressions, lighting circumstances, and pose variations impact the effectiveness of automatic age estimation [9].

The age classification of facial images has received a lot of attention in recent years. CNN in particular has gained a lot of academic attention because they can learn a compact and discriminative feature representation from large-scale data. As a result, CNN has had a lot of success in the community vision, dramatically improving classification performance [9].

Rather than estimating exact ages, this study focuses on the issue of age group classification. Although several age classification methods have been presented, most of them have focused on constrained images [9], such as Morph [10] and FG-NET [11]. Unlike the approaches described above, this study focuses on unconstrained images and investigates the problem of excessive class imbalance across various age groups: many face images belong to the 51-80 years old age group, while only a few belong to the Adolescence age group. This paper proposes a CNN model using standard techniques with FaceNet embeddings as the feature vectors [12]. The contributions of this study are outlined two folds. (1) The experimental results demonstrated that the suggested strategy may assist with the issue of class imbalance in CNN-based classification. Despite the extremely difficult nature of the images in the AAF datasets, our approach method is acceptable for face recognition and it is being utilized in this study to recognize age groups. (2) The proposed work attempted the tedious and complex task of classification of age into eight groups based on the books: The Human Odyssey [13] by splitting AAF dataset.

The contents of this paper are arranged as follows: Section 2 provides a quick review of the relevant work. Section 3 describes the proposed method in detail. Meanwhile, Section 4 contains the experimental setup. Results and discussions are in Section 5. Finally, conclusions are formed in Section 6.

2. Related Work
The hierarchical review organized in [14] provided a detailed study of age representation methods from face images. Bodhe et al. [15] presented a ScatterNet Inception Hybrid Network (SIHN) network for age-invariant face recognition that learns deep features. The SIHN network employs ScatterNet characteristics, allowing it to learn usable representations quickly with less labelled examples. SIHN proposed a detailed and dependable system for autonomous face tracking applications with an age range estimation technique incorporated. This system, however, can only distinguish a face when it is in a frontal position. Zhang et al. [16] provided a unique method for labelling a large number of in-the-wild facial images with high-quality posterior age distributions. A probability distribution of estimated ages for a face is provided by each posterior. Their technique is based on the fact that determining who is the older of two persons is easier than determining the person's actual age.

In automated age estimation, feature extraction is a critical stage. Certain methods for feature extraction have been presented, for example, Local Binary Pattern (LBP) [17], Active Appearance Model (AAM) [18], Biologically-Inspired Features (BIF) [19] and Anthropometric Features [20]. Regression and classification methods are also used to determine the precise age or age group from face images. In [21], Support Vector Machine (SVM) was used to classify people by their age. To accurately predict the age, regression techniques such as Support Vector Regression (SVR) [18], linear regression
[2], and Canonical Correlation Analysis (CCA) [22] were used. Despite the fact that all of these approaches were showed to be effective on constrained benchmarks, they may not be able to deal with large variations in unconstrained images [9].

CNN had recently investigated age estimation from face images, which attracted a lot of interest among researchers as depicted in table 1. Zhang et al. [23] suggested a novel Residual networks of Residual (RoR) networks architecture for age and gender classification in high-resolution face images in the wild. Smith and Chen [8] suggested employing deep CNN to use transfer learning to solve the challenge of distinguishing the age and gender of a person from an image. The Label Distribution Age Encoding (LDAE) employed in their study understands that people age differently, which aids in determining a person's age. Zhang et al. [24] offered a new approach called Recurrent Age Estimation (RAE) that takes use of both appearance and individual aging trends. RAE employs the Convolutional Neural Network Long-Short Term Memory (CNN-LSTM) architecture, with CNN being trained to extract discriminative appearance characteristics from face images and LSTM being used to learn individualized aging patterns using personal information sequences. For age-invariant face recognition, Shakeel and Lam [25] offered a robust deep-feature encoding-based discriminative model. Using a pre-trained deep-CNN model, they learn high-level deep features. Taheir and Toygar [26] offered a novel deep neural network architecture for age estimate called Directed Acyclic Graph Convolutional Neural Networks (DAG-CNNs), which takes advantage of multi-stage information from multiple layers of a CNN. For age estimation, Liu et al. [27] used a multi-task learning (MTL) network merging classification and regression called Classification Regression Multi-Task (CR-MT) network, where regression relies on classification as an auxiliary task. Besides that, when compared to multi-class classification approaches, an age ranking strategy based on CNN is employed in age estimation [28] to get reduced estimation errors.

Table 1. Research on age classification using CNN.

| Author and year       | Basic architecture | Dataset            |
|-----------------------|--------------------|--------------------|
| Zhang et al., 2017 [23]| ResNet             | IMDB-WIKI-101      |
|                       |                    | ImageNet           |
|                       |                    | Adience            |
| Smith and Chen, 2018 [8]| VGGNet            | Morph-II           |
| Zhang et al., 2019 [24]| LSTM              | Morph              |
|                       |                    | FG-NET             |
| Shakeel and Lam, 2019 [25]| AlexNet          | Morph              |
|                       |                    | FG-NET             |
|                       |                    | LAG                |
| Taheir and Toygar, 2019 [26]| VGG16 / GoogLeNet| Morph-II           |
| Liu et al., 2020 [27]| AlexNet            | Morph-II           |
|                       |                    | Webface            |
|                       |                    | CACD               |
|                       |                    | IMDB-WIKI          |
| Chen et al., 2017 [28]| LeNet              | Morph              |

To summarize, CNN had a lot of success with the age classification problem, greatly by enhancing classification performance. Most of the previous research improved classification accuracy by adjusting network design. Nevertheless, age group classification with an unbalanced data distribution has not yet been thoroughly examined. Therefore, this study presented a CNN model that is capable of classifying the age groups with unbalanced data distribution.
3. Proposed Method

This section explains in detail about the proposed work that consists of FaceNet that uses CNN to classify the age group for unbalanced data distribution.

FaceNet directly trains its output to be a compact 128-D embedding Large Margin Nearest Neighbour (LMNN) using a triplet-based loss function based on the huge margin nearest neighbour. The loss attempts to divide the negative and positive sets by a distance margin. The triplets consist of two matching face thumbnails, and the loss attempts to distinguish the positive from the negative pair by a distance margin. The thumbnails are close-ups of the face; no 2D or 3D alignment is done aside from size and translation. Selecting which triplets to utilize turns out to be critical for obtaining high performance and, as influenced by curricular learning [29].

Open Source Computer Vision (OpenCV) Library is a free and open-source software library for machine learning and computer vision. OpenCV was developed to offer a standard foundation for computer vision applications and to accelerate the adoption of machine perception into commercial goods. While OpenCV was utilized to assist with face recognition, it was not responsible for identifying faces. OpenCV can execute facial and accurate face detection by utilizing a pre-trained deep learning face detector model included with the library.

Two crucial deep learning phases must be applied to the OpenCV face recognition pipeline: 1) face detection spots the existence and position of a face in an image but does not determine it, and 2) to extract the 128-d feature vectors (also known as "embeddings") that quantify each face in the image. The model that quantifies each face in an image comes from the OpenFace project [30], a Python and Torch implementation of deep learning face recognition. This implementation is based on a 2015 CVPR publication by Schroff et al., [12].

The FaceNet implementation in the OpenCV facial recognition pipeline that is used for age group classification is shown in figure 1. The first step is input an image into the face recognition pipeline. Face detection is a technique for detecting the position of a face in an image. Facial landmarks can be computed as an option, allowing the face to be pre-processed and aligned. Face alignment is the process of determining the geometric structure of the faces and aiming to get a canonical alignment of the face based on rotation, scale, and translation, as the name suggests. Face alignment, while optional, has been shown to improve face recognition accuracy in various pipelines. After applying face alignment and cropping, run the input face through the deep neural network.

![Figure 1. Age group classification process.](image-url)
The most significant component of the methodology is the end-to-end learning of the entire system, given the model specifics and treating it as a black box (see figure 2). To that purpose, this study used a triplet loss that exactly matches what proposed work wants to achieve in terms of face classification.

Figure 2. Structure of the CNN model

The input data to the network and the triplet loss function are included in the training process. Although the planned work did not make a direct comparison to other losses, such as those involving pairs of positives and negatives, it did provide some suggestions. Each input batch of data for a deep learning face recognition model has three images: the anchor, the positive picture, and the negative picture (see figure 3). The argument is that the anchor and positive image are both from the same person or face, however the negative image does not. The anchor and positive image's 128-d embeddings are closer together. At the same time, the embeddings for the negative image are pushed further away. The network is able to learn to quantify faces in this way and deliver extremely robust and discriminating embeddings appropriate for face recognition as a result. The triplet loss, on the other hand, attempts to establish a buffer between each pair of faces from one person to all others. This allows one identity's faces to exist on a manifold while maintaining distance and consequently discriminability from other identities [12].

Figure 3. The Triplet Loss reduces the distance between an anchor and a positive with the same identity while increasing the distance between an anchor and a negative with a different identity.

4. Experimental Setup

The preliminary experiment is to classify the age group using CNN model. The specifications of the personal computer (PC) used are as follows; 4 GB RAM, central processing unit (CPU) Intel i7-3770 and Windows 10 Pro operating system. The age group classification model is implemented using the latest version of Python 3.9.0. Two experiments have been conducted for this study. The purpose of each experiment is to find out the performance of the CNN model for age group classification. Experiment 1 is conducted to validate the training accuracy while Experiment 2 is conducted to validate the testing accuracy. The number of test dataset is varied in both experiments. In this section, we introduce our data preparing and the evaluation matrix of the proposed method for age group classification.
4.1. Dataset

The dataset used in these experiments is AAF (All-Age Face) dataset [31] that is publicly available online. The AAF dataset contains a huge number of face images of individuals of all ages that were collected in unrestricted conditions. The dataset contains 13,322 face images (mostly Asian) distributed across all ages from 2 to 80 years, including 7,381 females and 5,941 males.

AAF dataset is separated into eight ranges of the human age as shown in table 2. The age group is used as a reference in [13]. The duplicate image is randomly discarded from the real AAF dataset during the data cleaning phase. The total of the AAF dataset after data cleaning is 13,316. Some examples of the images in the AAF dataset are demonstrated in figure 4.

| Age group               | Years old | Number of data |
|------------------------|-----------|----------------|
| Adolescence            | 12-20     | 1212           |
| Early Adulthood        | 21-35     | 5198           |
| Early Childhood        | 4-6       | 381            |
| Infancy                | 2-3       | 277            |
| Late Childhood         | 9-11      | 340            |
| Mature Adulthood       | 51-80     | 2306           |
| Middle Childhood       | 7-8       | 249            |
| Midlife                | 36-50     | 3353           |

Figure 4. Example of image in each age group, but the experiments are only conducted on two age groups which are Adolescence and Mature Adulthood.

In these experiments, only two types of age groups are used which are Adolescence and Mature Adulthood. The AAF dataset is quite large. The Adolescence and Mature Adulthood age groups were chosen because they have approximately the same number of data and were acceptable for comparison. Furthermore, the large age disparities between these two age groups facilitated in the categorization procedure for the purpose of the work.
In Experiment 1, the same dataset is used for training and testing which is 2463 or 70% from the overall dataset from Adolescence and Mature Adulthood the age group. While the dataset in Experiment 2 applied a holdout method by splitting data 70% for training (2463) and 30% for testing (1055). Both Experiment 1 and Experiment 2 used the same training dataset. Most of the images come from the Mature Adulthood age group. Table 3 and table 4 show data splitting in the preliminary experiment.

| Dataset                  | Age group       | Number of data |
|--------------------------|-----------------|----------------|
| Training data / Testing data | Adolescence    | 1614           |
|                          | Mature Adulthood| 849            |

| Dataset | Age group       | Number of data | Splitting of data |
|---------|-----------------|----------------|-------------------|
| Training data | Adolescence | 2463           | 1614              |
|          | Mature Adulthood | 849            | 70%               |
| Testing data    | Adolescence   | 1055           | 363               |
|          | Mature Adulthood | 692            | 30%               |

4.2. Evaluation matrix
Given that the objective is often to obtain movement data without using other techniques of observation, the proposed study states that training and testing dataset should be separated by individual rather than randomly.

The accuracy performance for both experiments is evaluated using a confusion matrix. A confusion matrix, also known as an error matrix or a contingency table, is a predetermined table structure that enables perception of the proposed model's efficiency [32]. The confusion matrix is used in the domain of machine learning approaches to demonstrate classification accuracy for target class and output class. A confusion matrix includes information about actual and expected grouping class. The data in the matrix is commonly used to evaluate the execution of certain frameworks. The disarray grid for a two-class classifier is shown in figure 5.

| Actual | Predicted | |
|--------|-----------|-----------|
|        | Positive (+) | Negative (-) |
| Positive (+) | True Positive (TP) | False Negative (FN) |
| Negative (-) | False Positive (FP) | True Negative (TN) |

- TP: The number of right expectations that an occurrence is certain
- FP: The number of inaccurate expectations that a case negative
- FN: The number of inaccurate expectations that a case is certain
- TN: The number of right expectations that an occurrence is negative
- P: The number of positive
- N: The number of negative

Figure 5. Confusion matrix binary classification
Besides calculating the accuracy, this study also considered the F1-score, sensitivity, specificity, precision, and false positive rate. The explanation of the confusion matrix attributes can be found in table 5.

**Table 5. Attributes in confusion matrix.**

| Measure               | Derivations                        | Explanation                                                                 |
|-----------------------|------------------------------------|-----------------------------------------------------------------------------|
| Accuracy (ACC)        | \((TP + TN) / (P + N)\)            | Accuracy is a metric that measures how many true predictions of the model produced throughout the whole test dataset. |
| F1-score (F1)         | \(2TP / (2TP + FP + FN)\)          | The greater the F-score, the more accurate the model. A model's accuracy decreases as its F-score decreases.       |
| Sensitivity (SN)      | \(TP / (TP + FN)\)                | Sensitivity or the true positive rate, is a measure of how many true positives are predicted out of all the positives in the dataset. |
| Specificity (SPC)     | \(TN / (FP + TN)\)                | Specificity is the percentage of true negatives that are accurately identified.                                      |
| Precision (PPV)       | \(TP / (TP + FP)\)                | Precision is a measure of the accuracy of a positive prediction.                                                           |
| False Positive Rate (FPR) | \(FP / (FP + TN)\)              | The false positive rate is a measure of how many positive findings are predicted out of all negative situations.          |

5. Result and Discussion

This section elaborates the result of two experiments. The outcome of Experiment 1 is shown in table 6. For Experiment 1, 70% of the overall dataset of Adolescence and Mature Adulthood age group is trained (2463/3518). Total correct rate of the Experiment 1 is 84.90% which means, from 2463 total test dataset, only 2091 is correctly labelled. Table 7 shows the result achieved in Experiment 2 that used hold out methods for data splitting. 70% from the overall data is the train dataset and another 30% is the test dataset. Total correct rate in Experiment 2 is increased to 0.22% from Experiment 1. That means the total number dataset that is correctly labelled is 898 over 1055 from the total test dataset.

**Table 6. Result for Experiment 1 by age group**

| Age group       | Sample size | Correctly labelled | Correct rate | Total correct rate / Accuracy |
|-----------------|-------------|--------------------|---------------|-------------------------------|
| Adolescence     | 849         | 636                | 74.91%        | 84.90% (2091/2463)            |
| Mature Adulthood| 1614        | 1455               | 90.15%        |                               |

**Table 7. Result for Experiment 2 by age group**

| Age group       | Sample size | Correctly labelled | Correct rate | Total correct rate / Accuracy |
|-----------------|-------------|--------------------|---------------|-------------------------------|
| Adolescence     | 363         | 215                | 59.22%        | 85.12% (898/1055)             |
| Mature Adulthood| 692         | 683                | 98.70%        |                               |
Figure 6 indicates the result of metrics derived from the confusion matrix in both experiments. In general, there are no significant changes between Experiments 1 and 2 in terms of accuracy and other attributes. The unbalanced dataset for both age group classification (Adolescence and Mature Adulthood) is used in the experiment that deteriorated the performance in terms of accuracy specifically.

| Actual | Experiment 1 | Experiment 2 |
|--------|--------------|--------------|
|        | Predicted    | Predicted    |
|        | Adolescence | Mature Adulthood | Adolescence | Mature Adulthood |
| Adolescence | 636 | 159 | 215 | 9 |
| Mature Adulthood | 213 | 1455 | 148 | 683 |

| Measure                  | Experiment 1 | Experiment 2 |
|--------------------------|--------------|--------------|
| Accuracy (ACC)            | 84.90%       | 85.12%       |
| F1-score (F1)             | 77.37%       | 73.25%       |
| Sensitivity (SN)          | 80.00%       | 95.98%       |
| Specificity (SPC)         | 87.23%       | 82.19%       |
| Precision (PPV)           | 74.91%       | 59.23%       |
| False Positive Rate (FPR) | 12.77%       | 17.81%       |

Figure 6. Result of metrics derived from confusion matrix in the experiment.

The accuracy, F1-score, sensitivity, specificity, precision and false positive rate of each experiment was recorded. Based on the result, it shows a significant decrease in the Adolescence age group dataset with just a little performance drop. It is possible to conclude that the dataset itself is unbalanced in both training and testing datasets, with Mature Adulthood images exceeding Adolescence. However, the experimental results show that CNN is still capable of achieving competitive classification accuracy for two age groups in the AAF dataset with unbalanced data distribution.

Furthermore, the result of this study has been compared with previous research [33] that used the same AAF dataset but applied a different method as shown in table 8. Gong et al. [33] proposed a de-biasing adversarial network (DeBrain) that determines to extract disentangled feature representations and indicating the methods used to tackle the problems of age classification. Based on the comparison, the result of this study shows a competitive classification accuracy using the proposed model although the accuracy value is slightly reduced.

| Experiment | Method            | Accuracy |
|------------|-------------------|----------|
| Experiment 1 | CNN (FaceNet model) | 84.90%   |
| Experiment 2 | CNN (FaceNet model) | 85.12%   |
| Gong et al. [36] | DebFace       | 94.45%   |

Table 8. Comparison of results with related work.

6. Conclusion
In this paper, CNN model is proposed to classify two age groups from AAF dataset with unbalanced data distribution: Adolescence and Mature Adulthood. The CNN model that is used in this study is entirely data-driven, learning its representation straight from the pixels of the face. Rather of employing
constructed features, this study used a large AAF dataset including the face images of persons in different ages from 2 years to 80 years. It seems obvious that age group classification models exhibit biased performance when compared to different age groups. Although the distribution of the dataset is unbalanced for two age groups in AAF dataset, the experimental results show that the CNN model is still capable of achieving competitive classification accuracy.

In the future the existing CNN model will be improved to achieve better performance in terms of accuracy for age group classification. The selection of the dataset using the balance class will be considered in future work for better accuracy results.

Acknowledgements
This work was supported by the Fundamental Research Grant Scheme (FRGS) under the Ministry of Education (MOE) with grant number 600-IRMI/FRGS 5/3 (234/2019).

References
[1] Razalli H and Alkawaz M H 2019 Real-Time Face Tracking Application with Embedded Facial Age Range Estimation Algorithm 9th International Conference on System Engineering and Technology vol 6 (IEEE) pp 2–7
[2] Htet K S and Myint Sein M 2020 Effective Marketing Analysis on Gender and Age Classification with Hyperparameter Tuning 2nd Global Conference on Life Sciences and Technologies (IEEE) pp 247–8
[3] Sangaralingam K, Verma N, Datta A, Chugh V and Ravi A 2018 Predicting Age & Gender of Mobile Users at Scale - A Distributed Machine Learning Approach International Conference on Big Data (IEEE) pp 1817–26
[4] Ferizal R, Wibirama S and Setiawan N A 2017 Gender Recognition using PCA and LDA with Improve Preprocessing and Classification Technique 7th International Annual Engineering Seminar (Yogyakarta)
[5] Manju D and Radha V 2020 A Novel Approach for Pose Invariant Face Recognition in Surveillance Videos International Conference on Computational Intelligence and Data Science vol 167 (Elsevier B.V.) pp 890–9
[6] Ahamed H, Alam I and Islam M 2018 HOG-CNN Based Real Time Face Recognition International Conference on Advancement in Electrical and Electronic Engineering (Gazipur: IEEE) pp 24–7
[7] Huang L, Guo J and Gao Z 2020 A Face Recognition System on Embedded Device J. Comput. 31 176–83
[8] Smith P and Chen C 2018 Transfer Learning with Deep CNNs for Gender Recognition and Age Estimation 5th National Symposium for NSF REU Research in Data Science, Systems, and Security (IEEE) pp 2564–71
[9] Liu W, Chen L and Chen Y 2018 Age Classification Using Convolutional Neural Networks with the Multi-class Focal Loss 3rd International Conference on Automation, Control and Robotics Engineering
[10] Nair A K U, Almadan A and Rattani A 2020 Understanding Fairness of Gender Classification Algorithms Across Gender-Race Groups 19th IEEE International Conference on Machine Learning And Applications
[11] Srinivas N, Atwal H, Rose D C, Mahalingam G, Ricanek K and Bolme D S 2017 Age, Gender, and Fine-Grained Ethnicity Prediction using Convolutional Neural Networks for the East Asian Face Dataset 12th International Conference on Automatic Face & Gesture Recognition (IEEE) pp 953–60
[12] Schroff F, Kalenichenko D and Philbin J 2015 FaceNet: A Unified Embedding for Face Recognition and Clustering Computer Society Conference on Computer Vision and Pattern Recognition (IEEE) pp 815–23
[13] Armstrong T 2019 *The Human Odyssey: Navigating the Twelve Stages of Life* (Mineola, New York: Ixia Press)

[14] Rwigema J, Mfitumukiza J and Tae-yong K 2021 A hybrid approach of neural networks for age and gender classification through decision fusion *Biomed. Signal Process. Control* 66 102459

[15] Bodhe S, Kapse P and Singh A 2019 Real-time Age-Invariant Face Recognition in Videos Using the ScatterNet Inception Hybrid Network (SIHN) *International Conference on Computer Vision Workshop* (IEEE/CVF) pp 1112–20

[16] Zhang Y, Liu L, Li C and Loy C C 2017 Quantifying Facial Age by Posterior of Age Comparisons *The British Machine Vision Conference*

[17] Salihbašic A and Orehovacki T 2019 Development of Android Application for Gender, Age and Face Recognition using OpenCV *42nd International Convention on Information and Communication Technology, Electronics and Microelectronics* pp 1635–40

[18] Choi S E, Jo J, Lee S, Choi H, Kim I J and Kim J 2017 Age face simulation using aging functions on global and local features with residual images *Expert Syst. Appl.* 80 pp 107–25

[19] Tian Q and Chen S 2018 Joint gender classification and age estimation by nearly orthogonalizing their semantic spaces *Image Vis. Comput.* 69 pp 9–21

[20] Duan M, Li K, Yang C and Li K 2018 A hybrid deep learning CNN–ELM for age and gender classification *Neurocomputing* 275 pp 448–61

[21] Rafique I, Asad M, Hamid A, Awais M, Naseer S and Yasir T 2019 Age and Gender Prediction using Deep Convolutional Neural Networks *International Conference on Innovative Computing* (IEEE)

[22] Boussaad L and Boucetta A 2020 An effective component-based age-invariant face recognition using Discriminant Correlation Analysis *J. King Saud Univ. - Comput. Inf. Sci.*

[23] Zhang K E, Gao C E, Guo L, Sun M, Yuan X, X. Han T, Zhao Z and Li B 2017 Age Group and Gender Estimation in the Wild With Deep RoR Architecture *The Chinese Conference on Computer Vision* vol 5 (IEEE Access) pp 22492–503

[24] Zhang H, Geng X, Zhang Y and Cheng F 2019 Recurrent age estimation *Pattern Recognit. Lett.* 125 pp 271–7

[25] M.S. Shakeel and K.-M. Lam 2019 Deep-feature encoding-based discriminative model for age-invariant face recognition *Pattern Recognit.* 93 pp 442–57

[26] Taheri S and Toygar Ö 2019 On the use of DAG-CNN architecture for age estimation with multi stage features fusion *Neurocomputing* 329 pp 300–10

[27] Liu N A, Zhang F A N and Duan F 2020 Facial Age Estimation Using a Multi-Task Network Combining Classification and Regression *IEEE Access* vol 8 (IEEE) pp 92441–51

[28] Chen S, Zhang C, Dong M, Le J and Rao M 2017 Using Ranking-CNN for Age Estimation *Conference on Computer Vision and Pattern Recognition* (IEEE) pp 742–51

[29] Huang Y, Wang Y, Tai Y, Liu X, Shen P, Li S, Li J and Huang F 2020 CurricularFace: Adaptive Curriculum Learning Loss for Deep Face Recognition *Conference on Computer Vision and Pattern Recognition*

[30] Amos B, Ludwiczuk B and Satyanarayanan M 2016 OpenFace : A general-purpose face recognition library with mobile applications pp 1–18

[31] Cheng J, Li Y, Wang J, Yu L and Wang S 2019 Exploiting Effective Facial Patches for Robust Gender Recognition *Tsinghua Science and Technology* 24 pp 333–45

[32] Masood S, Gupta S, Gupta S and Ahmad M 2018 Prediction of human ethnicity from facial images using neural networks *Data Engineering and Intelligent Computing*

[33] Gong S, Liu X and Jain A K 2020 Jointly De-biasing Face Recognition and Demographic Attribute Estimation *European Conference on Computer Vision* (Springer, Cham) pp 330–47