Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 11, No. 2, 2023, 229-236.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

A Deep Learning-based Model for Gender Recognition in Mobile Devices

Fatemeh Alinezhadand, Kourosh Kiani* and Razieh Rastgoo

Electrical and Computer Engineering Department, Semnan University, Semnan, Iran.

Article Info

Article History:
Received 09 December 2022
Revised 11 January 2023
Accepted 09 March 2023

DOI:10.22044/jadm.2023.12487.2398

Keywords:
Deep learning, Gender recognition, Margin distillation, Dense block, MobileFaceNet.

**Corresponding author: Kourosh.kiani@semnan.ac.ir (K. Kiani).*

Abstract

Gender recognition is an attractive research area in the recent years. To make a user-friendly application for gender recognition, having an accurate, fast, and lightweight model applicable in a mobile device is necessary. Although successful results have been obtained using the Convolutional Neural Network (CNN), this model needs high computational resources that are not appropriate for mobile and embedded applications. In order to overcome this challenge and considering the recent advances in deep learning, in this paper, we propose a deep learning-based model for gender recognition in mobile devices using the lightweight CNN models. In this way, a pretrained CNN model, entitled Multi-Task Convolutional Neural Network (MTCNN), is used for face detection. Furthermore, the MobileFaceNet model is modified and trained using the margin distillation cost function. To boost the model performance, the dense block and depthwise separable convolutions are used in the model. The results on six datasets confirm that the proposed model outperforms the MobileFaceNet model on six datasets with the relative accuracy improvements of 0.02%, 1.39%, 2.18%, 1.34%, 7.51%, 7.93% on the LFW, CPLFW, CFP-FP, VGG2-FP, UTKFace, and own data, respectively. In addition, we collect a dataset including a total of 100’000 face images from both male and female in different age categories. Images of the women are with and without headgear.

1. Introduction

With the advent of Artificial Intelligence (AI) as well as deep learning, visual understanding has become increasingly relevant to the computer vision society [1]. Furthermore, different research areas have benefited from the AI capabilities [1-9]. Due to the emergence of social platforms, age and gender classification have been around for quite some time now, and continual efforts have been made to improve its results. Using the biometric properties has an important role in many applications. For example, recently, the e-security aims to find some alternatives for the individual’s authentication and recognition. To this end, the human face can be used to illustrate many characteristics such as ethnicity, gender, age, and emotions. Furthermore, an accurate classification approach for detecting the aforementioned characteristics can improve the performance of many applications. In this paper, we only concentrate on one of these characteristics, gender, aiming to propose a model for gender classification from the input image.

Emerging trend in developing the accurate and handy applications embedded in intelligent mobile devices has attracted many attentions in different research areas. To make a user-friendly application, having an accurate, fast, and lightweight model applicable in a mobile device is necessary. Though successful results have been achieved using CNN [4, 6], this model needs high computational resources that are not appropriate for mobile and embedded applications. To cope this challenge, some small and lightweight CNN models have been designed such as MobileNetV1 [2], ShuffleNet [10], and MobileNetV2 [11]. Considering these models, we propose a CNN-based model for gender classification using the dense block and margin distillation approaches. The proposed model successfully performs on mobile devices. The results on six datasets confirm the classification accuracy improvement compared state-of-the-art baseline model in the field. Our main contributions can be listed as follows:

- **Model:** We proposed a CNN-based model using dense block and margin distillation approaches. The proposed model successfully performs on mobile devices. The MobileFaceNet model is modified and trained using the margin distillation cost function. The model performance is boosted by benefiting from the dense block and depthwise separable convolutions.
- **Mobile API:** We developed an API for gender classification on mobile devices.
- **Dataset:** We collected a dataset including a total of 100'000 face images from both males and females in different age categories. Images of the women are with and without headgear.
- **Performance:** The proposed model is evaluated on six datasets. The results confirm that the proposed model has a competitive performance compared to the baseline models.

The rest of this paper is organized as what follows. Section 2 presents a brief review on the recent works in the face and gender classification. The proposed model is introduced in details in Section 3. The results are reported in Section 4. Finally, we discuss and conclude on the work in Section 5.

2. Related Works

We briefly review the recent works in gender classification. Different features can be used in this research area. In some research works, the features obtained from the fingerprints are used for gender classification. In this way, different methods can be used for feature extraction from the fingerprints such as Discrete Wavelet Transform (DWT) [12], Singular Value Decomposition (SVD) [13], and Local Binary Pattern (LBP) [14]. Gnanasivam and Muttan suggested a model for gender classification using the DWT, SVD, and K Nearest Neighbor (KNN). In this way, the energy is calculated from all the sub-bands of DWT and combined with the spatial features of non-zero singular values achieved from the SVD of fingerprint images. The evaluation results on a dataset including 3570 fingerprints of 1980 male fingerprints and 1590 female fingerprints showed the overall classification accuracy of 88.28% [15]. Gornale *et al*. used three categories of local texture features of the fingerprint images for gender classification. These categories include the image quality, Local Binary Patterns (LBP), and the extracted features. The model evaluation has led to the classification accuracy of 95.88% for this model [16]. Kruti *et al*. studied the impact of feature level fusion and synthesis of classifiers for gender recognition using fingerprints. After feature extraction from a single instance of fingerprint, the feature level fusion and synthesis of classifiers on fingerprint are done. The results on four datasets confirm the efficiency of the feature level fusion with synthesis of classifiers compared to the non-fused and single classifier [17].

In some works, the facial analytics are used for real-time gender classification. In this way, the facial features are extracted using different models, such as Convolutional Neural Network (CNN). Undru *et al*. used the pre-trained model, VGG16, combined with the Haar Cascade classifier to classify the images based on the facial traits. A recognition accuracy of 90% on a dataset including 6000 images achieved with this method [18]. Zeni *et al*. developed a real-time end-to-end gender detector using Deep Neural Networks (DNNs). The proposed model performs a gender classification on the images with various pose, illumination, and occlusions. In this way, some annotations have been added to the Pascal VOC 2007 and CelebA datasets to train the model. The experimental results confirm the efficiency of the additional annotations in the model, obtaining a better performance for gender classification [19]. Arriaga *et al*. designed a CNN-based model for real-time gender classification on a constrained platform. To this end, a real-time vision system has been developed to simultaneously recognize the face, gender, and emotion. The accuracy of 96% and 66% have been reported on the IMDB and FER-2013 emotion dataset, respectively [20].

In addition to the aforementioned features, the features obtained from the face can be used for gender classification. Different methods can be used for feature extraction from the face images such as deep learning-based models, especially the CNN models. Smith *et al*. designed a deep learning-based model using the pre-trained CNN models such as VGG19 and VGGFace for gender classification. They performed different analysis on the effects of changes in various design schemes and training parameters in the model. Furthermore, the separate models are used for male and female recognition. The results on the MORPH-II dataset show a gender recognition accuracy of 98.7% [21].

Choudhary *et al*. developed a framework using the combination of a CNN and a KNN-classifier for gender classification from an input image. They analyzed the impact of layer number on the recognition accuracy of the model, concluding the increasing model accuracy with increasing the layers number [22]. Listio studied the comparison of two CNN-based models, Inception-V3 and MobileNet, for gender recognition using eye images. The results on a public dataset including a total of 2,681 images consisting of 1251 male eyes and 1430 female eyes shows that the Inception-V3 model is better than the MobileNet in gender recognition, obtaining a recognition accuracy of 91.82% [23]. Janahiraman and Subramaniam have made a comparison of three CNN-based models, VGG16, ResNet-50, and MobileNet, for gender recognition. The evaluation results have been reported on a dataset including 1000 images of Asian faces inclusive of Malaysians and some Caucasians, confirming the superiority of VGG-16 model compared to other models using the recognition accuracy metric [24]. Khalifa *et al.* designed a CNN-based model for gender recognition using the iris segmentation from a background image. In their model, 16 layers including three convolutional layers with different convolution window sizes as well as three fully connected layers have been used. The results on a dataset containing of 3000 images have led to the recognition accuracy of 98.88% [25]. Savchenko proposed a two-stage model for age/gender recognition as well as face identification. The MobileNet is modified and used for face recognition. The detected faces are grouped using hierarchical agglomerative clustering techniques. Finally, gesture recognition is performed using the aggregation of predictions for individual images. The results on the LFW dataset confirmed the recognition accuracy of 94.1% [26].

3. Proposed Model

Here, we present the details of the proposed model. Also an overview of the model can be found in Figure 1.

3.1. Face detection

The pre-trained model, Multi-Task Convolutional Neural Network (MTCNN) [27], is used for face detection and alignment (Figure 2). MTCNN is a three-stage cascaded framework that initially resizes the input image into different scales and constructs an image pyramid. A fully convolutional network, namely Proposal Network (P-Net), is employed to obtain the bounding box regression vectors corresponding to the candidate windows. After that, the Non-Maximum Suppression (NMS) method is applied to candidate windows and unify highly overlapped candidates. Another CNN, entitled Refine Network (R-Net), is used to choose only the relevant candidate windows.

3.2. Alignment and cropping

The Ms-celeb-1m dataset [28] is used as model training. The detected face from MTCNN is aligned and cropped to 112 × 112 shape. Some samples of this process have been shown in Figure3.

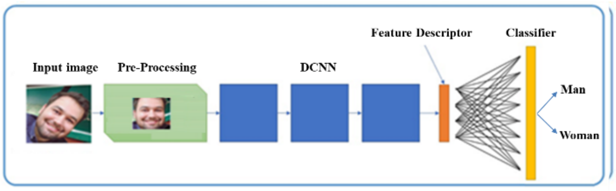

Figure 1. An overview on the proposed model.

Figure 3. Face detection step in the proposed model.

3.3. Model training

The MobileFaceNet [2] with some modifications is used in the proposed model. In the next sub-section, details of the model architecture will be presented. The margin distillation cost function is used for model training that the teacher network is a ResNet100 model [29] trained using ArcFace

Figure 4. Pre-processing step of the proposed model including Alignment and cropping.

[30] cost function, as shows in the following:

$$L_{ArcFace} = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s\left(\cos\left(\theta_{y_i}+m\right)\right)}}{e^{s\left(\cos\left(\theta_{y_i}+m\right)\right)}+\sum_{j=1,j\neq y_i}^{n} e^{scod\theta_j}} \quad (1)$$

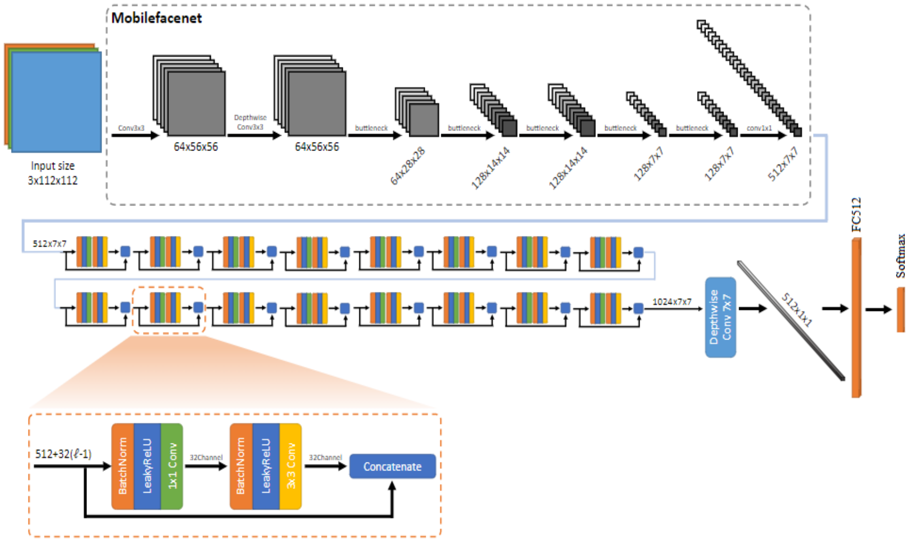

Figure 5. Proposed architecture of the proposed model including the Dense Block, Depth-wise convolution layer, and FC.

where y_i is the class label corresponding to the deep feature of the i-th sample, x_i . N and n are the batch size and the class number, respectively. The embedding feature dimension is set to 512 in this paper. θ_j is the angle between the j-th weight and the i-th feature. The embedding feature x_i is fixed by L2 normalization and re-scaled to s. m is an additive angular margin penalty between x_i and the weight, aiming to simultaneously enhance the intra-class compactness and inter-class discrepancy.

3.4. Model architecture

We used the MobileFaceNet [2] with some modifications. In this way, the last two layers of the MobileFaceNet model are removed, and the model weights of the convolution layers are preserved. The last convolution layer contains 512 feature maps with 7 x 7 shape. Furthermore, a dense block containing 16 dense layers is stacked on top of the model. The input feature maps obtained from the last convolution layer of the MobileFaceNet model are convolved using a 1 x 1 convolution in the first dense layer. After that, a 3 x 3 convolution is applied to 32 feature maps achieved from the 1 x 1 convolution. All of the feature maps of 32 channels are concatenated to feed to the next dense layer. This process is repeated for 16 dense layers of the dense block. In addition to this, a depth-wise convolution layer is stacked on the dense block. A 512 7 x 7 depth-wise convolution is used to get the 512 x 1 x 1 feature vector. Finally, a Fully Connected (FC) layer is used on top of the previous convolution layer. Details of the dense block including 16 dense layers, a depth-wise convolution layer, and a FC layer are shown in Figure 4.

3.5. API implementation

Using Java programming language and the Android SDK [31] can help the developers to implement the code in a managed and controllable environment. In addition to benefiting from the Java language and Android platform, we also use the ncnn framework as a neural network-based framework in mobile devices to run faster than the other frameworks on mobile devices [32]. Relying on the efficient NCNN implementation, the developers can easily deploy deep learning algorithm models to the mobile platform. To make a relation between the ncnn in C++ and Java programming language, NDK is used. The proposed API works with the Android 7 or higher versions. The proposed API can be found in Figure 5. This API includes two main functionalities as follows:

- **Add person:** In this functionality, we can add a person to the API. At least three images from different viewpoints for each person are needed here.

- **Gender recognition:** This functionality performs the recognition task by using the proposed CNN-model. To this end, the API asks the person to take an image. This image is processed to recognize the gender of the person in this image.

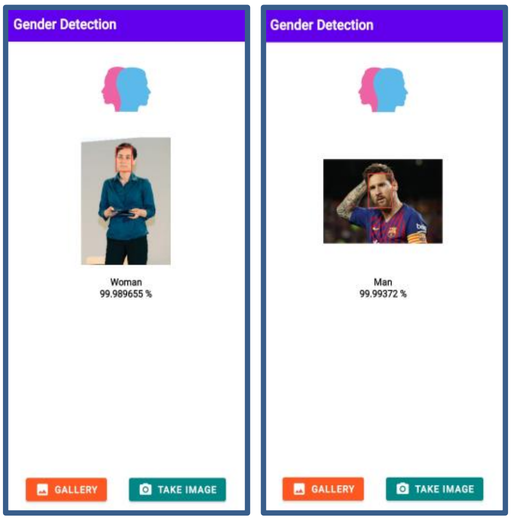

Figure 6. Proposed API.

4. Results

In this section, we present the details of the achieved results for the proposed model.

4.1. Implementation details

The results of the model evaluation have been obtained using an Intel(R) Xeon(R) CPU E5-2699 (2 processors) with 90GB RAM, Python software, and the MXNet library on a NVIDIA GPU. The AdaGrad and learning rate of 0.005 have been used. The proposed model is trained for total of 2000 epochs by early stopping. All of the results are reported on the test data.

4.2. Datasets

Seven datasets have been used for model training and evaluation. Here, we briefly review the details of these datasets (Table 4 and Figure 6, 7, and 8).

- **Labeled Faces in the Wild (LFW) [33]:** There are 13233 images from 5749 subjects with limited variations in pose, age, expression, and illuminations in this dataset.
- **Cross-Pose Labeled Faces in the Wild (CPLFW) [34]:** This dataset is a modified version of the LFW dataset including the pose difference of the same person with preserving the identities of the LFW at the same time.
- **CFP [35]:** This dataset includes the celebrity images in frontal and profile views. We use only the CFP-Frontal-Profile (CFP_FP) evaluation protocol of this dataset.
- **VGG2_FP [36]:** The dataset contains a total of 3.31 million images of 9131 subjects, with an average of 362.6 images for each subject.
- **UTKFace [37]:** This dataset includes the 20,000 face images with long age span (range from 0 to 116 years old).
- **Own data:** We collected a dataset including a total of 100’000 face images from both males and females in different age categories. Images of the women are with and without headgear.

Figure 7. Some samples of LFW dataset.

Figure 7. Some samples of the CFP_FP dataset.

Figure 8. Some samples of the proposed dataset.

4.3. Model results

Here, we compare the proposed model with the MobileFaceNet, as the most similar API to our model. As one can see in Table 1, the proposed model outperforms the MobileFaceNet model on six datasets with the relative accuracy improvements of 0.02%, 1.39%, 2.18%, 1.34%, 7.51%, 7.93% on the LFW, CPLFW, CFP-FP, VGG2-FP, UTKFace, and own data, respectively.

Table 1. Comparison of results of the proposed model with the MobileFaceNet model.

Dataset	Model		
	MobileFaceNet	Proposed model	Relative improvement
LFW	99.50	**99.52**	0.02
CPLFW	86.63	**88.02**	1.39
CFP_FP	89.73	**91.91**	2.18
VGG2_FP	90.30	**91.64**	1.34
UTKFace	88.12	**95.63**	7.51
Own data	92.01	**99.94**	7.93

Furthermore, the operation, size, and time complexity of the proposed model compared with the MobileFaceNet has been shown in Table 2. Finally, the average recognition accuracies from 50 runs have been shown in Table 3. The proposed model obtained the average recognition accuracy of 99.52%, 88.02%, 91.91%, 91.64%, 95.63%, and 99.94% from 10 runs with standard deviation of 0.03, 0.04, 0.04, 0.05, 0.03, and 0.01 on the LFW, CPLFW, CFP_FP, VGG2_FP, UTKFace, and Own data datasets, respectively. In addition, the one-sample t-test analysis at a 5% alpha level has been performed for each dataset to compare the average accuracies of two runs (first run contains 50 times of run and the second one contains 10 times of run). As Table 3 shows, it is highly likely that the mean recognition accuracies of two runs are the same. In overall, the proposed model has promising results in both meal and female categories.

Table 2. Operation, size, and time complexity of the proposed model compared with the MobileFaceNet.

Complexity	Model	
	MobileFaceNet	Proposed model
FLOPs/10^8	4.39	4.92
Size / Mb	3.92	7.06
Time / ms	40.70	53.76

4. Conclusion and Future Work

In this paper, we proposed a deep learning-based model for gender recognition in mobile devices using the lightweight CNN models. We used the MTCNN model for face detection. In addition, the MobileFaceNet model was modified and trained using the margin distillation cost function. Also the dense block and depth-wise separable convolutions were used in the model to enhance the model performance. A dataset including a total of 100’000 face images from both males and females in different age categories has been collected. The results on six datasets confirmed that the proposed model outperformed the MobileFaceNet model with the relative accuracy improvements of 0.02%, 1.39%, 2.18%, 1.34%, 7.51%, 7.93% on the LFW, CPLFW, CFP-FP, VGG2-FP, UTKFace, and own data, respectively. In a future work, we aim to use the vision transformer for gesture recognition from input images/videos.

Table 3. Statistical analysis of the proposed model on six datasets (p-value = 0.05).

Dataset	Statistical analysis	
LFW	H_0	$\mu = 99.52\%$
	H_1	$\mu \neq 99.52\%$
	t-test	H_0 is accepted
CPLFW	H_0	$\mu = 88.02\%$
	H_1	$\mu \neq 88.02\%$
	t-test	H_0 is accepted
CFP_FP	H_0	$\mu = 91.91\%$
	H_1	$\mu \neq 91.91\%$
	t-test	H_0 is accepted
VGG2_FP	H_0	$\mu = 91.64\%$
	H_1	$\mu \neq 91.64\%$
	t-test	H_0 is accepted
UTKFace	H_0	$\mu = 95.63\%$
	H_1	$\mu \neq 95.63\%$
	t-test	H_0 is accepted
Own data	H_0	$\mu = 99.94\%$
	H_1	$\mu \neq 99.94\%$
	t-test	H_0 is accepted

References

[1] B. Patel and Sh. Degadawala, “A Survey Paper on Gender Classification using Deep Learning”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 6, 2020. DOI: https://doi.org/10.32628/CSEIT20613.

[2] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand et al. “Mobilenets: Efficient convolutional neural networks for mobile vision applications”. arxiv:1704.04861, 2017.

[3] R. Rastgoo and V. Sattari Naeini, “A neuro-fuzzy QoS-aware routing protocol for smart grids”, in *2014 22nd Iranian Conference on Electrical Engineering (ICEE)*, Tehran, Iran, pp. 1080-1084, 2014.

[4] R. Rastgoo and V. Sattari Naeini, “Tuning parameters of the QoS-aware routing protocol for smart grids using genetic algorithm”, Applied Artificial Intelligence, vol. 30, no. 1, pp. 52-76, 216.

[5] R. Rastgoo, K. Kiani, and S. Escalera, “Sign language recognition: A deep survey”, Expert Systems with Applications, vol. 164, p. 113794, 2021.

[6] B. Patel and Sh. Degadawala, “A Survey Paper on Gender Classification using Deep Learning”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 6, 2020. DOI: https://doi.org/10.32628/CSEIT20613.

[7] R. Rastgoo and V. Sattari Naeini, “Gsomcr: Multi-constraint genetic-optimized QoS-aware routing protocol for smart grids”, in *Iranian Journal of Science and Technology, Transactions of Electrical Engineering,* Tehran, Iran, vol. 42, pp. 185-194, 2018.

[8] N. Majidi, K. Kiani, and R. Rastgoo, “A deep model for super-resolution enhancement from a single image”, Journal of AI and Data Mining, vol. 8, no. 4, pp. 451-460, 2020.

[9] K. Kiani, R. Hematpour, and R. Rastgoo, “Automatic Grayscale Image Colorization using a Deep Hybrid Model”, Journal of AI and Data Mining, vol. 9, no. 3, pp. 321-328, 2021.

[10] X. Zhang, X. Zhou, M. Lin, and J. Sun, “Shufflenet: An extremely efficient convolutional neural network for mobile devices”. arXiv:1707.01083, 2017.

[11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks”, *in Conference on Computer Vision and Pattern Recognition*, Salt Lake City,, USA, 2018.

[12]https://mil.ufl.edu/nechyba/www/eel6562/course_materials/t5.wavelets/intro_dwt.pdf. Access data: Feb 3, 2023.

[13]https://www.cs.cmu.edu/~venkatg/teaching/CStheory-infoage/book-chapter-4.pdf. Access data: Feb 3, 2023.

[14] J. Lizé, V. Débordès, H. Lu, K. Kpalma, and J. Ronsin, “Local binary pattern and its variants: application to face analysis”, in *First International Conference on Smart Information and Communication Technologies,* Saidia, Morocco, pp.94-102, 2020.

[15] P. Gnanasivam and S. Muttan, “Fingerprint Gender Classification using Wavelet Transform and Singular Value Decomposition”, International Journal of Computer Science, vol. 9, no. 2, 2012.

[16] S.S. Gornale, M. Basavanna, and R. Kruthi, “Fingerprint Based Gender Classification Using Local Binary Pattern”, International Journal of Computational Intelligence Research, vol. 13, no. 2, pp. 261-271, 2017.

[17] R. Kruti, A. Patil, and Sh. Gornale, “Fusion of Features and Synthesis Classifiers for Gender Classification using Fingerprints”, International Journal of Computer Sciences and Engineering, vol. 7, no. 5, 2019.

[18] T.R. Undru and C. Anuradha, “A Real Time Gender Recognition System Using Facial Images and CNN”, International Journal of Computer Sciences and Engineering, vol. 7, no. 9, 2019.

[19] L.F. Zeni and C. Jung, “Real-Time Gender Detection in the Wild Using Deep Neural Networks”, in *31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI),* Brazil, 2018.

[20] O. Arriaga, P.G. Pl¨oger, and M. Valdenegro, “Real-time Convolutional Neural Networks for Emotion and Gender Classification”, arXiv:1710.07557, 2017.

[21] Ph. Smith and C. Chen, “Transfer Learning with Deep CNNs for Gender Recognition and Age Estimation”, in *IEEE International Conference on Big Data (Big Data),* Seattle, WA, USA, 2018.

[22] S. Choudhary, M. Agarwal, and M. Jailia, “Design Framework for Facial Gender Recognition Using MCNN”, International Journal of Engineering and Advanced Technology (IJEAT), vol. 8, no. 3, 2019.

[23] S.W.P. Listio, “Performance of Deep Learning Inception Model and MobileNet Model on Gender Prediction Through Eye Image”, Sinkron: Jurnal dan Penelitian Teknik Informatika, vol. 7, no. 4, 2022.

[24] T.V. Janahiraman and P. Subramaniam, “Gender Classification Based on Asian Faces using Deep Learning”, in *IEEE 9th International Conference on System Engineering and Technology (ICSET),* Shah Alam, Malaysia, 2019.

[25] N. M. Khalifa, M.H. N. Taha, A.E. Hassanien, and H. N. E. T. Mohamed, “Deep Iris: Deep Learning for Gender Classification Through Iris Patterns”, ACTA INFORM MED, vol. 27, Issue 2, pp. 96-102, 2019. doi: 10.5455/aim.2019.27.96-102.

[26] A. V. Savchenko, “Efficient facial representations for age, gender and identity recognition in organizing photo albums using multi-output ConvNet”, PeerJ Comput. Sci, pp. 1-26, 2019. doi:10.7717/peerj-cs.197.

[27] K. Zhang, Z. Zhang, Z. Li, Y. Qiao, “Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks”, IEEE Signal Proc. Let., vol. 23, no. 10, pp. 1499–1503, 2016.

[28] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, “Ms-celeb-1m: A dataset and benchmark for large-scale face recognition”, arXiv: 1607.08221, 2016.

[29] K. He, X. Zhang, Sh. Ren, and J. Sun, “Deep Residual Learning for Image Recognition”, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Las Vegas, NV, USA, 2016.

[30] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, “ArcFace: Additive Angular Margin Loss for Deep Face Recognition”, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Long Beach, CA, USA, 2019.

[31] https://google-developer-training.github.io/android-developer-fundamentals-course-concepts-v2/unit-1-get-started/lesson-1-build-your-first-app/1-0-c-introduction-to-android/1-0-c-introduction-to-android.html. Access date: Feb 3, 2023.

[32] https://github.com/Tencent/ncnn, Access date: Feb 3, 2023.

[33] http://vis-www.cs.umass.edu/lfw/, Access date: Feb 3, 2023.

[34] T. Zheng and W. Deng, “Cross-Pose LFW: A Database for Studying Cross-Pose Face Recognition in Unconstrained Environments”, Beijing University of Posts and Telecommunications, Technical Report, 2018.

[35] S. Sengupta, J.Ch. Chen, C. Domingo, V.M. Patel, R. Chellapp, and D.W. Jacobs, “Frontal to profile face verification in the wild”, in *IEEE Winter Conference on Applications of Computer Vision, WACV* 2016, Lake Placid, NY, USA, 2016.

[36] Q.Cao, L. Shen, W. Xie, O.M. Parkhi, and A. Zisserman, “VGGFace2: A dataset for recognizing faces across pose and age”, arXiv:1710.08092, 2018.

[37]https://www.kaggle.com/datasets/jangedoo/utkface-new. Access date: Feb 3, 2023.

یک مدل مبتنی بر یادگیری عمیق برای تشخیص جنسیت در دستگاه‌های تلفن همراه

فاطمه علی‌نژاد، کوروش کیانی * و راضیه راستگو

دانشکده برق و کامپیوتر، دانشگاه سمنان، سمنان، ایران.

ارسال ۲۰۲۲/۱۲/۰۹؛ بازنگری ۲۰۲۳/۰۱/۱۱؛ پذیرش ۲۰۲۳/۰۳/۰۹

چکیده:

تشخیص جنسیت یک حوزه تحقیقاتی جذاب در سال‌های اخیر است. با توجه به پیشرفت‌های اخیر در یادگیری عمیق، در این مقاله، ما یک مدل مبتنی بر یادگیری عمیق برای تشخیص جنسیت در دستگاه‌های تلفن همراه پیشنهاد می‌کنیم. برای این منظور، یک شبکه عصبی کانولوشن (CNN) از پیش آموزش دیده، با عنوان شبکه عصبی کانولوشن چند وظیفه‌ای (MTCNN)، برای تشخیص چهره استفاده می‌شود. علاوه بر این، مدل MobileFaceNet با استفاده از تابع هزینه تقطیر حاشیه، اصلاح و آموزش داده شده است. برای افزایش عملکرد مدل، از کانولوشن های جداشدنی Dense Block و Depthwise در مدل استفاده شده است. نتایج بر روی شش مجموعه داده تأیید می‌کند که مدل پیشنهادی با بهبود دقت نسبی ٪۰،۰۲، ٪۱،۳۹، ٪۲،۱۸، ٪۱،۳۴، ٪۷،۵۱، ٪۷،۹۳ بر روی پایگاه داده‌های LFW، CPLFW، CFP-FP، VGG2-FP، UTKFace و داده‌های خود از مدل MobileFaceNet بهتر است. علاوه بر این، ما مجموعه داده‌ای را جمع آوری کردیم که در مجموع، صد هزار تصویر چهره از زن و مرد در رده‌های سنی مختلف را شامل می‌شود. تصاویر زنان با پوشش سر و بدون پوشش سر می‌باشد.

کلمات کلیدی: یادگیری عمیق، تشخیص جنسیت، تقطیر دانش، بلوک متراکم (Dense block)، MobileFaceNet.