An Efficient Face Gender Detector on a CPU with Multi-Perspective Convolution

Adri Priadana, Muhamad Dwisnanto Putro and Kang-Hyun Jo
Department of Electrical, Electronic, and Computer Engineering
University of Ulsan
Ulsan, Korea
Email: priadana3202@mail.ulsan.ac.kr, dputro@mail.ulsan.ac.kr, acejo@ulsan.ac.kr

***Abstract*—Intelligent digital signage demands software to work lightly in real-time on embedded devices. Smart digital signage requires gender detection to deliver relevant and targeted ads automatically. This study presents an efficient real-time face gender detector (FGenCPU) that can run smoothly on a CPU implemented in digital signage to decide immediately and provide relevant and targeted ads while the audience faces it. The proposed architecture contains Multi-Perspective Convolution (MPConvNet) consisting of two main modules: backbone and recognition. An efficient extractor module with a multi-perspective contextual track through the various kernel sizes is utilized to apprehend different feature areas of the object. This architecture employs few filters in every convolution layer and generates few parameters. It makes the detector operate at high speed for real-time. The training is conducted on the UTKFace dataset to generate efficiently weighted models. The MPConvNet achieves competitive performance with other common and light architectures on UTKFace, Adience, and LFW datasets. Furthermore, the detector can run 38 frames per second when performing on a CPU in real-time.**

Keywords—face gender, real-time detector, multi perspective convolution, digital signage, user demographic

I. Introduction

Due to the recent evolution of digital content and its technologies, advertising media patterns continually transform. One of the products of this development is the appearance of digital signage that has become one of the essential platforms for offline advertising in modern cities [1]. Digital signage has become progressively widespread and can be found in public areas [2], [3]. Compared to conventional posters and billboard promotions, advertising by using digital signage are more flexible to customize. Audience demographics (e.g., gender) are vital to effectively segment the audience and determine relevant ads for each audience [4]. Recognizing the audience's gander will create more targeted advertisements [5].

Conceptually, digital signage can recognize gender by detecting and analyzing the audience's face when facing it using a camera located on the monitor. With the rapid development of computer vision, real-time face-based gender recognition techniques supporting digital signage systems are investigated [6]–[8]. Real-time face detection and gender recognition in digital signage often encounter various problems, such as limited hardware specifications. The recognition technique must be able to be implemented on a CPU or low-cost device with acceptable accuracy to eliminate the dependency on expensive hardware.

The Convolutional Neural Network (CNN) is widely used for object detection and recognition as a deep learning technique. CNN has been proven to have outstanding classification performance to recognize gender [9]–[11]. Various CNN models were developed to build face-based gender recognition. Ranjan et al. [12] proposed a CNN model called HyperFace-ResNet that was made on the ResNet-101 architecture with 44,7 million number parameters to build a gender detector. The model united the geometrically wealthy features. It was obtained from the lower layers and semantically robust features from the deeper layers of the ResNet. Therefore, multi-task learning can leverage their linkage. It also incorporated the features using hierarchical element-wise addition. This model achieved 98% and 94% accuracy on the CelebA and LFWA datasets. Garain et al. [13] also developed a CNN model called GRA_Net that builds on the ResNet with Gated Residual Attention with 33 million number parameters to build a gender detector. It improved the Residual Attention Network by using a modified version of the Attention block's output. This model achieved 99.2% and 81.4% accuracy on UTKFace (In-the-wild Faces) and Adience dataset, respectively. The number of parameters in a CNN model will undoubtedly affect the training and inference stage speed. The fewer the number of parameters will improve the efficiency of the detector. Therefore, a CNN model with a relatively small number of parameters will support the ability of the real-time detector if applied to low-cost devices or CPUs.

Although it has not produced high speed for real-time gender recognition, CNN has also been implemented in digital signage. A CNN model based on MobileNetV2 was presented to develop real-time gender recognition implemented on digital signage [6]. It achieved an accuracy of almost 95% and could process 5 FPS on an ARM-based CPU. It used an intelligent camera installed above the monitor that performed real-time gender recognition and a component that dynamically changed the content projected on the screen based on the audience's gender. Greco et al. [14] have also presented a real-time face-based gender recognition in digital signage using CNN. It allows boosting the effectiveness of the advertisement campaigns. In this case, it can substitute the static contents shown on the screen with some dynamic advertisements. Therefore, advertisements can be customized automatically depending on the gender of the audience viewing the monitor. This study proposed a CNN model based

978-89-93215-23-6/22/$31.00 ACA

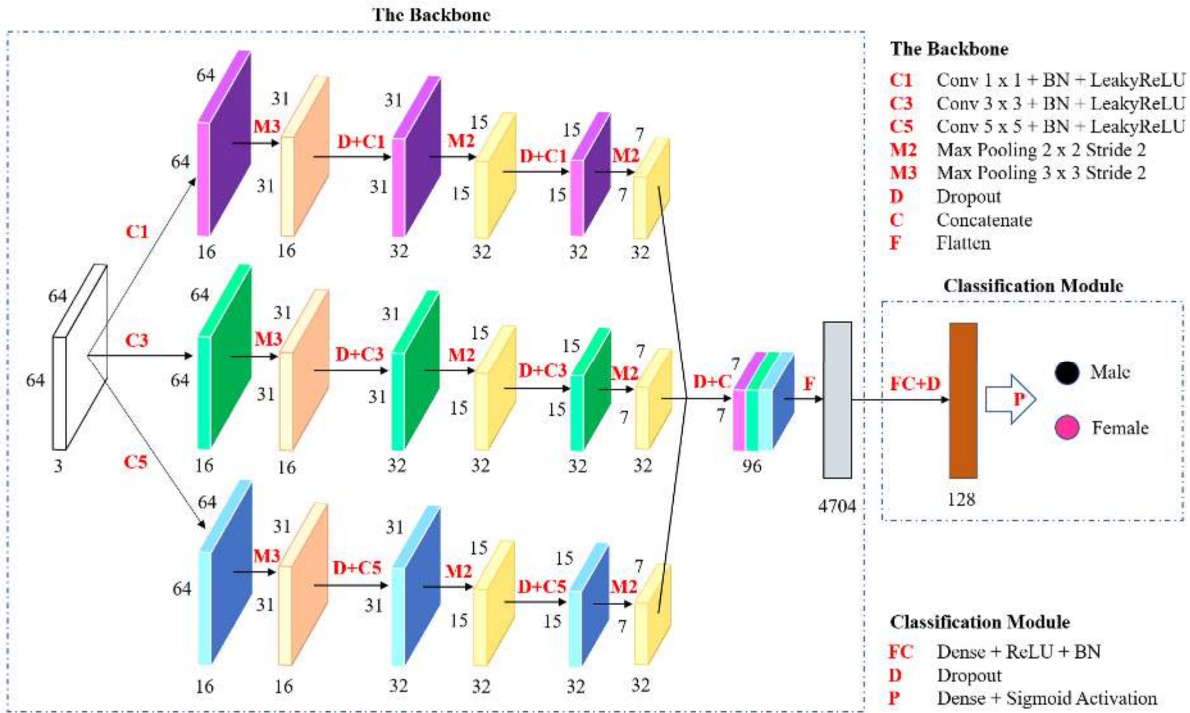

Fig. 1. The MPConvNet architecture for gender recognition.

on MobileNet with 96 × 96 inputs with three channels (RGB) that could process 3 FPS on a Raspberry Pi 3 Model B. However, the model still achieved 84,48% accuracy on the Adience dataset, although implemented based on a mobile or tiny version of the CNN model with 3,5 million number parameters. Therefore, a face-based gender detector with few parameters must be required to work smoothly and efficiently on digital signage used on a low-cost device or CPU-based. This study presents a new real-time detector that can efficiently detect a gender from frames containing a face.

An efficient real-time face gender detection (FGenCPU) proposes a light architecture with Multi-Perspective Convolution (MPConvNet) that offers a multi-perspective contextual track with various kernel sizes to deal with different sizes of the feature area of the object. The efficient means that this architecture uses few filters in every convolution layer and produces few parameters to make the detector run faster. Therefore, this architecture constructs a lightweight face gender detector that can swiftly implement on a low-cost device or CPU-based. The contribution of this study can summarize as follow:

1) A novel multi-perspective convolution (MPConvNet) is offered to extract features with various kernel sizes that run separately during the feature extraction part. It can capture information from different feature areas of the object.

2) An efficient convolution layer with a few filters is proposed to reduce the number of parameters. It can run smoothly on a CPU that is proper to implement for digital signage. The performance result achieves competitive accuracy with other architectures on UTKFace [15], Adience Benchmark [16], and Labeled Faces in the Wild (LFW) [17] datasets.

II. Proposed Architecture

The proposed architecture for gender recognition utilizes a series of convolution layers consisting of three perspective modules, as shown in Fig. 1. A backbone efficiently extracts face features by applying convolution with various kernel sizes that run separately. Moreover, the classification module is applied to predict the gender of the face. This architecture of gender recognition generated 659,650 parameters.

A. The Backbone of Gender Recognition

The backbone module extracts interest features from a face area. The features of an object will determine the level of performance of detection. Every object, especially a face, always has various feature area sizes. Each area's size represents a different perspective that needs to extract as a basis of prediction. The FGenCPU proposes a multi-perspective convolution MPConvNet by separating the convolution layer with a kernel's varying sizes that run parallel from the input layer to the end of the backbone module. This architecture consists of three perspective convolutions with 1 × 1, 3 × 3, and 5 × 5 kernel size. This variation aims to apprehend essential features from different sizes (small, medium, large) of the feature area and enrich the information of spatial features.

This technique is inspired by the Inception block model [18] that implemented different kernel sizes in the convolutions layer. In the MPConvNet architecture, these three perspectives are maintained separately from the beginning to the end of the extraction process. Every perspective consists of 3 convolution layers with 16, 32, and 32 filters. This perspective variation is put together at the end of the extraction process using concatenate operation. This approach keeps these perspectives independent as a packet feature map representing each perspective. Therefore, this perspective variation can be used as the main extractor for the gender recognition process in the classification module. This backbone accepts 64 × 64 inputs with three channels (RGB).

In order to manage the gradient problem, it applies batch normalization and Leaky ReLU (Leaky Rectified Linear Unit) activation after convolution operations. It also uses dropout before convolution layer on second and third layer to prevent overfitting [19], [20]. After these operations, the max-pooling layer with two strides also applies to shrink the feature map. This architecture does not apply 1 × 1 convolution before the expensive 3 × 3 and 5 × 5 convolution to calculate reductions because it will reduce the feature information that has been obtained. It uses a few filters in every convolution layer to redeem it. It provides an efficient architecture that supports the ability of the detector to operate smoothly and quickly on a low-computing device or CPU. As the last layer after concatenating, flatten operation is applied to make a one-dimensional vector so that it can be connected with the fully connected layer on the classification module.

B. Classification Module

The classification module is required on the detector to predict the gender of the face at the network's end. This module consists of a fully connected layer and an activation function. In this architecture, the fully connected layer has 128 units. It generates a few trainable parameters to reduce computation costs. After a fully connected layer, ReLU (Rectified Linear Unit) activation and Batch Normalization are applied to manage the gradient problem. It also uses dropout after these operations to prevent overfitting. The activation function transforms the real values generated in the previous layer to possibilities and always returns a value between 0 and 1 that represent the output as a male or female. The sigmoid function is described as:

$$S(x) = \frac{1}{1+e^{-x}} \quad (1)$$

where x is a logic score from the neural network and e is Euler's number.

C. Face Detector

This study performs face detection formerly before computing the facial gender prediction. It employed LWFCPU [21] that has good performance when applied to a CPU in real-time. It proceeds to obtain the face area called Region of Interest (RoI). The LWFCPU consists of twelve convolutional layers and uses six types of anchors that generate lightweight parameters. Additionally, it can operate quickly in real-time on CPU or low-computing devices. Furthermore, the RoI of the face is represented by bounding boxes, and this area will be cropped and scaled in accordance with the size of the gender recognition input.

III. Implementation Setup

The model is training using Google Colab with a GPU accelerator setting. The training is conducted and evaluated on UTKFace datasets. It trains with 10^{-2} as the learning rate and a total of 100 epochs. It uses Adam as an optimizer to update the weight to minimize error for Binary Cross-Entropy loss and 32 batch size to allow computational speedups from the parallelism of GPUs. Then, it is tested on Intel Core I5-6600 CPU @ 3.30GHz, 32GB RAM. The whole structure of MPCovNet is implemented on the Tensorflow 2.0 and Keras 2.3.1 framework.

TABLE I. Evaluation results on UTKFace, Adience, and LFW datasets

Architecture	Number of Parameter	Validation Accuracy (%)
Evaluation on UTKFace (Aligned and Cropped Faces)		
Krishnan et al. (VGG-16) [10]	138,357,544	91.90
Krishnan et al. (VGG-19) [10]	143,667,240	91.50
Krishnan et al. (ResNet-50) [10]	25,636,712	91.60
Savchenko (MobileNetV1) [11]	3,491,521	91.95
Hamdi & Moussaoui (Manually-designed) [22]	530,034	89.97
VGG-13 + Batch Normalization	34,467,906	90.36
VGG-11 + Batch Normalization	34,413,698	90.61
ResNet50	23,591,810	88.48
ResNet50V2	23,568,898	88.17
InceptionV3	21,806,882	87.35
MobileNet V1	3,230,914	89.82
MobileNet V2	2,260,546	89.71
Squeezenet + Batch Normalization	735,306	89.27
MPConvNet (Manually-designed)	**659,650**	**92.32**
Evaluation on Adience Benchmark		
Opu et al. (Manually-designed) [23]	210,050	85.77
Greco et al. (MobileNetsV1) [14]	3,538,984	84.48
Althnian et al. (VGG-16 without Fully Connected Layer) [24]	15,473,190	83.30
MPConvNet (Manually-designed)	**659,650**	**85.67**
Evaluation on LFW		
Greco et al. (MobileNetsV1) [14]	3,538,984	98.73
Althnian et al. (VGG-16 without Fully Connected Layer) [24]	15,473,190	72.50
FaceHop (Manually-designed) [25]	16,900	94.63
MPConvNet (Manually-designed)	**659,650**	**96.30**

IV. Experimental Results

The performance of the gender prediction is evaluated based on some datasets. This section investigates the proposed architecture on a benchmark consisting of UTKFace, Adience, and LFW datasets. It also evaluates the efficiency of the detector tested on a CPU and compares it to other competitors.

A. Evaluation on Datasets

1) UTKFace (Aligned and Cropped Faces): The proposed detector is evaluated in the UTKFace dataset to test the performance of the person detector on cropped faces. There are 23,708 face images on this dataset labeled in age, gender, and ethnicity. The face images ranging in age from 0 to 116 are included in this dataset. The images cover enormous variations such as pose, illumination, resolution, expression, etc. The dataset also presents aligned and cropped faces format. This study uses the aligned and cropped face image version dataset. In this experiment, the dataset is split into 70% and 30% as training and testing sets with a random permutation split. For instance, the distribution of 23,708 images is 16,596 for the train and 7,112 for the test. As a result, MPConvNet with 659,650 parameters achieves 92,32% Validation Accuracy (VA), exceeding common

(a)

(b)

Fig. 2. The correct prediction results (a), and the incorrect prediction results of the FGenCPU detector (b).

architecture VGG, ResNet, Inception, and MobileNet. Moreover, MPConvNet achieves a validation accuracy outperforming light architecture Squeezenet and Hamdi & Moussaoui, which differed by 3.05 and 2.35, respectively. Therefore, MPConvNet can work more effectively in gender recognition to support digital signage in real-time.

2) Adience Benchmark: The Adience dataset consists of 26,580 face images from 2,284 unique subjects with age and gender labels. The face images ranging in age from 0 to more than 60 are included in this dataset. The images cover enormous variations such as pose, noise, appearance, lighting, and more. The dataset is available in some formats. This experiment also does some pre-processing on the dataset, such as removing data that contains missing values. The aligned images dataset is used and split into 70% training sets and 30% testing sets with a random permutation split. For instance, the distribution of 17,492 images is 12,244 for the train and 5,248 for the test. As a result, MPConvNet with 659,650 parameters achieves 85.67 VA, exceeding common architecture VGG-16 and MobileNetsV1. In addition, the proposed detector is below the performance of Opu et al., which differed by 0.1 with 210,050 parameters, as can be seen in Table I.

3) Labeled Faces in the Wild (LFW): The LFW dataset consists of 13234 face images from 5,749 unique subjects collected from the web, with a significant imbalance between females (23%) and males (77%). This dataset is only labeled in the name of the person captured. In this study, the dataset is manually labeled related to gender. In this experiment, the dataset is split into 70% training sets and 30% testing sets with a random permutation split. The result generates 9,263 images for the train and 3,971 images for the test. As a result, MPConvNet with 659,650 parameters achieves 96.30 VA, exceeding common architecture VGG-16 and light architecture FaceHop. In addition, the proposed detector is below the performance of MobileNetsV1 with 3.5 million parameters, which differed by 2.43, as can be seen in Table I.

B. Runtime Efficiency

The MPConvNet is specifically designed to be applied to low-cost devices such as digital signage. This lightweight architecture certainly supports the performance of the detection and recognition process. The proposed face gender detector generates 659,650 parameters which use an efficient number of kernels and computations. It promotes this architecture to perform in real-time on a low-cost device or CPU-based. The proposed detector achieves 47.08 FPS for

TABLE II. COMPARISON OF ARCHITECTURE SPEEDS ON CPU

Architecture	Gender (FPS)	Face + Gender (FPS)
VGG-16 + Batch Normalization	23.05	20.57
VGG-13 + Batch Normalization	27.46	24.57
VGG-11 + Batch Normalization	29.04	25.84
ResNet50	24.71	22.01
ResNet50V2	24.56	22.28
InceptionV3	23.37	21.25
MobileNet V1	38.49	32.73
MobileNet V2	33.12	28.56
Squeezenet + Batch Normalization	41.91	35.13
MPConvNet	**47.08**	**38.72**

gender recognition (Gender), and 38.72 FPS for the integrated with face detection using LWFCPU and proposed gender recognition (Face + Gender). Therefore, MPConvNet can work more efficiently in gender recognition to support digital signage in real-time. The proposed architecture becomes the fastest detector on a CPU compared to other common and light architecture that can be shown in Table II. Fig. 2 (a) shows the correct prediction results of the FGenCPU detector on the CPU. The red bounding box indicates a female face, and the green color indicates a male face.

C. Limitations

The FGenCPU detector with MPConvNet architecture is training on the UTKFace dataset to generate efficiently weighted models. The dataset covers enormous variations in the pose. Nevertheless, it doesn't have many instances for every variation. Furthermore, the pose variation of the face is no more than 30 degrees. Therefore, the FGenCPU detector, in some cases, has an incorrect prediction when it predicts the gender with face turned sideways fully and its variation. Fig. 2 (b) shows the incorrect prediction results of the FGenCPU detector. The FGenCPU detector also has satisfactory real-time detection accuracy for predicting the gender with the front face pose. It occurs when the audience faces the digital signage screen.

V. CONCLUSION

This paper presents an efficient real-time face gender detector that proposed a light architecture with Multi-Perspective Convolution (MPConvNet) to enrich information from the receptive field. The MPConvNet architecture is designed to perform smoothly in real-time on a CPU without compromising performance. The architecture produces a few parameters to shorten the operation time of the detector. The MPConvNet achieves competitive performance with other common and light architectures on the UTKFace, Adience Benchmark, and LFW datasets. The architecture also generates fewer parameters than the competitors on the Adience dataset with only utilizes 64 × 64 inputs with three channels (RGB). The detector can run 38.72 frames per second to detect face and recognize gender when working in real-time on a CPU. This result is faster than other competitors' architecture. In future work, the sharing of information between perspective techniques can explore in the backbone to increase the accuracy of the classification module.

ACKNOWLEDGMENT

This work was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (MOE)(2021RIS-003).

REFERENCES

[1] X. Zhang *et al.*, "A Digital Signage Audience Classification Model Based on the Huff Model and Backpropagation Neural Network," *IEEE Access*, vol. 8, pp. 71708–71720, 2020, doi: 10.1109/ACCESS.2020.2987717.

[2] X. Zhang *et al.*, "Analysis of Spatial Characteristics of Digital Signage in Beijing with Multi-Source Data," *ISPRS Int. J. Geo-Information 2019, Vol. 8, Page 207*, vol. 8, no. 5, p. 207, May 2019, doi: 10.3390/IJGI8050207.

[3] P. Morozova and N. Shilov, "Towards providing relevant digital signage advertisement to a group of users based on users' interests investigation," *Conf. Open Innov. Assoc. Fruct*, vol. 2019-April, pp. 263–268, May 2019, doi: 10.23919/FRUCT.2019.8711896.

[4] T. A. Tuan, T. D. Cao, and T. Truong-Huu, "DIRAC: A hybrid approach to customer demographics analysis for advertising campaigns," *Proc. - 2019 6th NAFOSTED Conf. Inf. Comput. Sci. NICS 2019*, pp. 256–261, Dec. 2019, doi: 10.1109/NICS48868.2019.9023806.

[5] A. Priadana, M. R. Maarif, and M. Habibi, "Gender Prediction for Instagram User Profiling using Deep Learning," in *2020 International Conference on Decision Aid Sciences and Application (DASA)*, 2020, pp. 432–436, doi: 10.1109/DASA51403.2020.9317143.

[6] A. Greco, A. Saggese, and M. Vento, "Digital Signage by Real-Time Gender Recognition from Face Images," *2020 IEEE Int. Work. Metrol. Ind. 4.0 IoT, MetroInd 4.0 IoT 2020 - Proc.*, pp. 309–313, Jun. 2020, doi: 10.1109/METROIND4.0IOT48571.2020.9138194.

[7] N. Turov, N. Shilov, and N. Teslya, "Digital signage personalization through analysis of the visual information about viewers," *Conf. Open Innov. Assoc. Fruct*, vol. 2019-April, pp. 444–450, May 2019, doi: 10.23919/FRUCT.2019.8711893.

[8] R. E. Abraham and M. Robert Kennedy, "Intelligent Digital Signage System Based on Gender Identification," *Lect. Notes Electr. Eng.*, vol. 492, pp. 251–262, 2018, doi: 10.1007/978-981-10-8575-8_25.

[9] P. Majumdar, R. Singh, and M. Vatsa, "Attention Aware Debiasing for Unbiased Model Prediction." pp. 4133–4141, 2021.

[10] A. Krishnan, A. Almadan, and A. Rattani, "Understanding Fairness of Gender Classification Algorithms across Gender-Race Groups," *Proc. - 19th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2020*, pp. 1028–1035, Dec. 2020, doi: 10.1109/ICMLA51294.2020.00167.

[11] A. V. Savchenko, "Efficient facial representations for age, gender and identity recognition in organizing photo albums using multi-output ConvNet," *PeerJ Comput. Sci.*, vol. 2019, no. 6, p. e197, Jun. 2019, doi: 10.7717/PEERJ-CS.197/TABLE-10.

[12] R. Ranjan, V. M. Patel, and R. Chellappa, "HyperFace: A Deep Multi-Task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 1, pp. 121–135, Jan. 2019, doi: 10.1109/TPAMI.2017.2781233.

[13] A. Garain, B. Ray, P. K. Singh, A. Ahmadian, N. Senu, and R. Sarkar, "GRA_Net: A Deep Learning Model for Classification of

Age and Gender from Facial Images," *IEEE Access*, vol. 9, pp. 85672–85689, 2021, doi: 10.1109/ACCESS.2021.3085971.

[14] A. Greco, A. Saggese, M. Vento, and V. Vigilante, "A Convolutional Neural Network for Gender Recognition Optimizing the Accuracy/Speed Tradeoff," *IEEE Access*, vol. 8, pp. 130771–130781, 2020, doi: 10.1109/ACCESS.2020.3008793.

[15] Z. Zhang, Y. Song, and H. Qi, "Age Progression/Regression by Conditional Adversarial Autoencoder." pp. 5810–5818, 2017.

[16] E. Eidinger, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces," *IEEE Trans. Inf. Forensics Secur.*, vol. 9, no. 12, pp. 2170–2179, Dec. 2014, doi: 10.1109/TIFS.2014.2359646.

[17] G. B. Huang, M. Mattar, T. Berg, E. Learned-Miller, and E. Learned-Miller, "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments," in *Preceding Workshop Faces Real-Life Images, Detection, Alignment, Recognition*, 2008, pp. 1–15.

[18] C. Szegedy *et al.*, "Going deeper with convolutions," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, pp. 1–9, 2015, doi: 10.1109/CVPR.2015.7298594.

[19] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 56, pp. 1929–1958, 2014.

[20] J. Liang and R. Liu, "Stacked denoising autoencoder and dropout together to prevent overfitting in deep neural network," *Proc. - 2015 8th Int. Congr. Image Signal Process. CISP 2015*, pp. 697–701, Feb. 2016, doi: 10.1109/CISP.2015.7407967.

[21] M. D. Putro, D. L. Nguyen, and K. H. Jo, "Lightweight convolutional neural network for real-time face detector on CPU supporting interaction of service robot," *Int. Conf. Hum. Syst. Interact. HSI*, vol. 2020-June, pp. 94–99, Jun. 2020, doi: 10.1109/HSI49210.2020.9142636.

[22] S. Hamdi and A. Moussaoui, "Comparative study between machine and deep learning methods for age, gender and ethnicity identification," *ISIA 2020 - Proceedings, 4th Int. Symp. Informatics its Appl.*, Dec. 2020, doi: 10.1109/ISIA51297.2020.9416549.

[23] M. Nahidul Islam Opu, T. K. Koly, A. Das, and A. Dey, "A Lightweight Deep Convolutional Neural Network Model for Real-Time Age and Gender Prediction," *Proc. 2020 3rd Int. Conf. Adv. Electron. Comput. Commun. ICAECC 2020*, Dec. 2020, doi: 10.1109/ICAECC50550.2020.9339503.

[24] A. Althnian, N. Aloboud, N. Alkharashi, F. Alduwaish, M. Alrshoud, and H. Kurdi, "Face Gender Recognition in the Wild: An Extensive Performance Comparison of Deep-Learned, Hand-Crafted, and Fused Features with Deep and Traditional Models," *Appl. Sci. 2021, Vol. 11, Page 89*, vol. 11, no. 1, p. 89, Dec. 2020, doi: 10.3390/APP11010089.

[25] M. Rouhsedaghat, Y. Wang, X. Ge, S. Hu, S. You, and C. C. J. Kuo, "FaceHop: A Light-Weight Low-Resolution Face Gender Classification Method," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12668 LNCS, pp. 169–183, Jan. 2021, doi: 10.1007/978-3-030-68793-9_12.