Pre-trained Convolutional Neural Networks for Gender Classification

Bhuvaneshwari Patil[1,2] and Mallikarjun Hangarge[3](✉)

[1] Gulbarga University, Kalaburagi, India
[2] Faculty at Presidency University, Bangalore, India
[3] Department of Computer Science, KASC College, Bidar, India
bsp4052001@gmail.com

Abstract. Many researchers have used Convolutional Neural Networks (CNN) models to solve the gender classification problem using pre-trained architectures. In this paper, the author has focused on investigating the success of the custom CNN model with respect to pre-trained deep neural models like VGG16, ResNet152V2, InceptionResNetV2 and EfficientNetV2L with limited data for gender classification.

Keywords: gender classification · keras models · convolution neural network · deep neural network

1 Introduction

Gender classification using facial features has received a lot of attention, but nowadays eye feature-based classification has gained attention among researchers. Feature extraction and classification procedures are crucial for implementing an effective automatic classification system. The traditional machine learning algorithms will give better accuracy, it is necessary to extract accurate features from eye datasets. On the other hand, Deep learning models will automatically extract features from raw data. Deep neural networks can investigate hidden and unpredictable feature sets, which can improve classification performance by exploring hidden and unpredictable feature sets when compared to typical machine learning.

Face traits gathered were from the rich study on human authentication features. In recent years, texture feature extraction (Majumdar & Patil, 2013) from an iris image has gained popularity as a soft biometric trait for determining a person's gender. When combined with related biometric data, the main benefit of employing soft biometrics is that it aids in the faster retrieval of identities by reducing the searching period. Iris data has been utilized productively in a wide range of settings, including airport check-in and refugee control (Bobeldyk & Ross, 2019). It can also be used in cross-spectral matching scenarios (Dantcheva et al., 2016) when comparing RGB and NRI images. By enhancing recognition qualities and accuracy, more semantic information about an unknown situation will be provided, filling the gap between machine and human descriptions of the area.

R. Manza et al. (Eds.): ACVAIT 2022, AISR 176, pp. 319–326, 2023.
https://doi.org/10.2991/978-94-6463-196-8_25

The area of interest has been narrowed to Eye region instead from whole face for Gender classification which reduces the computation complexity and provides the reliable results as the rate of change in eye features is negligible as compared with the changes in facial features. As a result, the dataset and the area of interest have been limited. Keras Applications, deep learning models that come with pre-trained weights and can be used to make predictions, extract features, and fine-tune them. These models can be built according to the required image dataset. Here, we are using VGG16, ResNet152V2, InceptionResNetV2, EfficientNetV2L models for eye images to classify as male and female eyes. The rest of the paper is organized as related work in the field of gender prediction/classification and deep neural network models in Section 2. Section 3 describes the applied methodology and dataset. Section 4 shows results and discussion of results. Finally, completed with conclusion.

2 Related Work

The first publication on gender prediction from geometric and texture aspects of iris pictures was published by Thomas et al. (Thomas et al., 2007). The author considered SVM and NN for classification with an accuracy of 80%. For an improved gender prediction rate, Tapia & Aravena (J. Tapia & Aravena, 2018) presented a modified Lenet-5 CNN model. Four convolution layers and one fully- connected layer with a minimum number of neurons make up the updated network. Bobeldyk and Ross (Bobeldyk & Ross, 2019).

Sreya & Jones (C & Jones B, 2020) investigated the IITD Dataset and employed ANN to recognise iris patterns. The writers went over each phase of the recognition process in great detail. To find the pupil region, the trials were undertaken out on cropped NIR pictures. The accuracy of prediction is said to be dependent on processing, according to the authors.

Singh et al. (Singh et al., 2018) used a variant of an auto-encoder that includes the attribute class label in addition to the reconstruction layer. They used images of NIR Oculars that had been scaled down to 48X64 pixels. For their method, they employed the GFI and ND-Iris-0405 Datasets. The researchers used RDF and NNet classifiers and got an accuracy of 83.17 percent. They claim that The Deep Class-Encoder takes a fifth of the total training time, and that their results exceed Tapia et al.'s (J. E. Tapia et al., 2016).

For an improved gender prediction rate, Tapia & Aravena (J. Tapia & Aravena, 2018) developed a modified Lenet-5 CNN model. Four convolution layers and one fully-connected layer with a small number of neurons make up the redesigned network. To avoid the risk of over-fitting and solve the two-class gender prediction problem, a minimal number of neurons is recommended. The authors used Data Augmentation to boost the size of each eye's Dataset from 1,500 to 9,500 images. The authors find that combining CNN for the right and left eyes results in better prediction than each eye alone.

Deep learning has been used for gender classification from facial images, drawing inspiration from a variety of fields. The study (Janahiraman et al., 2019) develops a dataset consisting of facial photographs of Caucasians and Malaysians, and then applies several Convolutional Neural Networks for gender prediction. Using the VGG-16, ResNet-50, and MobileNet models, it reports accuracy of 88 percent, 85 percent, and 49 percent, respectively. On the Adience dataset (Eidinger et al., 2014), (Akbulut et al., 2017) applied CNN and LRA-ELM methods and achieved 80% and 87.13%, respectively. In the study (Abdalrady et al., 2020), typical CNN models were swapped out for the PCANet model for gender categorization. Furthermore, it is possible to reduce the size of the network architecture in intricate CNN models by employing PCANet.

3 Methodology

Several face datasets have been used for gender classification in the literature, including Adience, FERET (color-feret-database, 2021), Gallagher's dataset (Gallagher et al., 2008), and LFW (Huang et al., 2007). However, they all supply images of the entire face, and segmenting the eyes is an extra task for the researchers. We utilize the dataset labelled "Female and Male" which is referenced in this article since it conducts a comparison analysis between state-of-the-art CNN models simply utilizing eye pictures (eyes-rtte, 2021). The dataset consists of 5202 female and 6323 male eye images.

3.1 Keras Models

The input images are cropped eye portion from UKTFace images and resized to 75 × 75. The model is trained with 8068 images and tested using 3467 images. The dataset consist of 6323 male images and 5202 female images and the parameters used are as listed in Table 1.

The authors of this study want to understand how well different pre-trained deep learning models perform when it comes to gender classification from eye images. Here, we used VGG16, ResNet152V2, InceptionResNetV2, EfficientNetV2L models for this purpose and used dropout layer to protect the network against an overfitting issue. The 'Female and Male' dataset is split in the ratio of 70:30 as training and testing data.

The model's size, accuracy, number of parameters, depth and time inference steps for CPU and GPU as given in Table 2 (Keras applications, 2021). EfficientNetV2L model has a huge number of parameters however the maximum top-1 and top-5 accuracy.

Table 1. Training parameters of CNN models

Parameters	Values
Optimizer	Adam
Loss	Categorical_crossentropy
Shuffle	True
Number of epochs	30
batch_size	32

Table 2. Keras models

Model	Size (MB)	Parameters	Top-1 Accuracy	Top-5 Accuracy	Depth	Time (ms)/inference step (CPU)	Time (ms)/inference step (GPU)
VGG16	528	14,764,866	71.3%	90.1%	16	69.5	4.2
ResNet152V2	232	58,532,354	78.0%	94.2%	307	107.5	6.6
InceptionResNetV2	215	54,413,538	80.3%	95.3%	449	130.2	10.0
EfficientNetV2L	479	117,872,290	85.7%	97.5%	-	-	-

Table 3. Summary of the Custom CNN model

Model: "Sequential"

Layer (type)	Output Shape	Param #
Conv2d_203 (conv2D)	(None, 32, 32, 32)	896
Max_pooling2d_7 (MaxPooling2D)	(None, 16, 16, 32)	0
Conv2d_204 (conv2D)	(None, 16, 16, 16)	9248
Max_pooling2d_8 (MaxPooling2D)	(None, 8, 8, 32)	0
Flatten_4 (Flatten)	(None, 2048)	0
Dense_4 (Dense)	(None, 128)	262272
Dense_5 (Dense)	(None, 1)	129

Total params: 272,545
Trainable params : 272,545
Non-trainable params : 0

3.2 Custom CNN

We have developed a custom CNN to train and test the 'Female and Male' dataset for gender classification. The model is trained using 272,545 parameters as listed in summary Table 3. The model consists of two convolution layers, two maxpooling layers and two dense layers. The modelcheckpoint callback function used to save the weights when there is an improvement during training.

4 Results and Discussion

The authors of this project want to see if they can utilize pre-trained deep networks like VGG16, ResNet152V2, InceptionResNetV2, EfficientNetV2L to classify gender from eye images. In addition, custom CNN model trained with same set of parameters and results are as shown in Figs 1, 2, 3, 4 and 5. The accuracy for these models are tabulated in Table 4.

The Figures 1, 2, 3, 4 and 5 shows that VGG16, ResNet152V2 have similar accuracy for both training and validation data whereas custom CNN has smooth curve for training as compare to validation. It is observed that all the five models, learning rate decreased when the models are trained for more than thirty epochs.

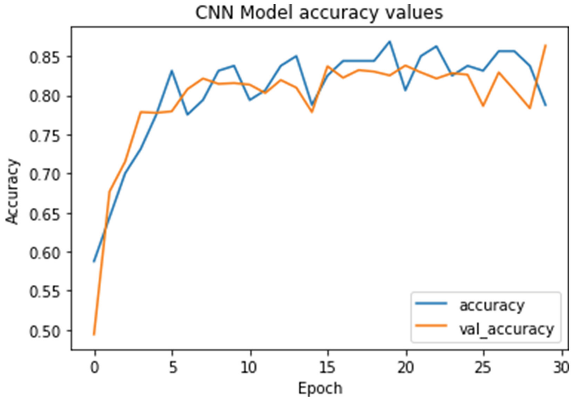

Fig. 1. Accuracy plot for VGG16

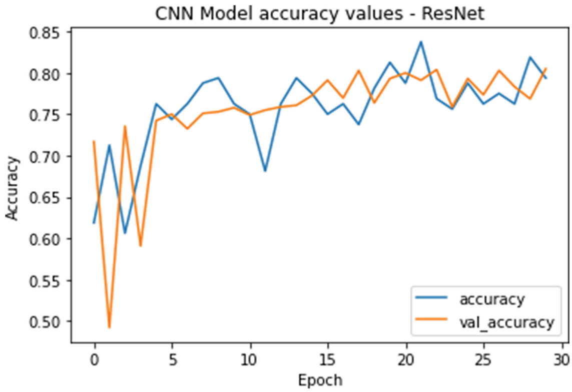

Fig. 2. Accuracy plot for ResNet152V2

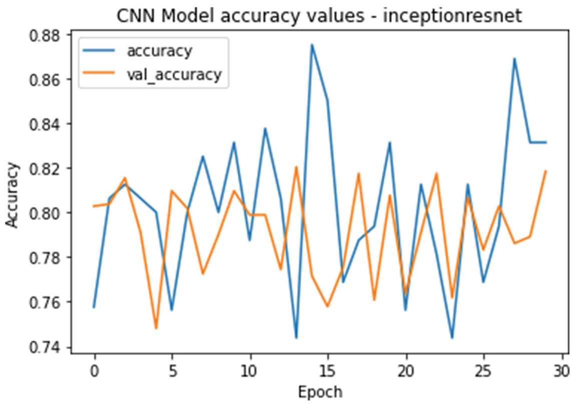

Fig. 3. Accuracy plot for InceptionResnetV2

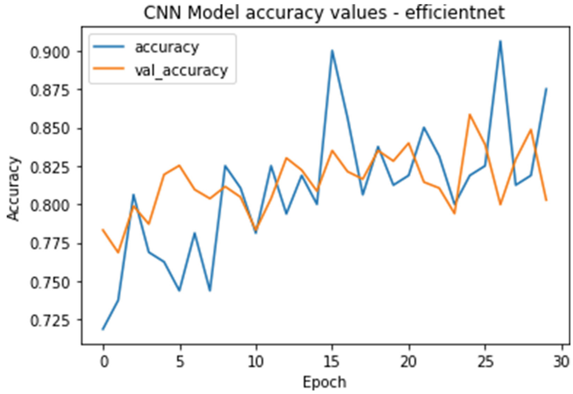

Fig. 4. Accuracy plot for EfficientNetV2L

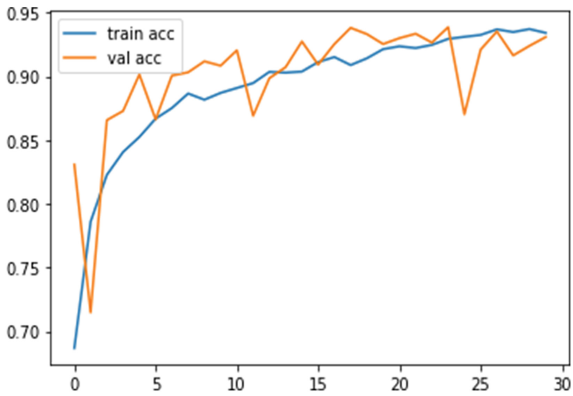

Fig. 5. Accuracy plot for custom CNN

Table 4. Gender Classification results

Model	Trainable parameters	Training duration (h:m:s)	Accuracy (%)
VGG16	50,178	0:37:40.917232	84.05
ResNet152V2	200,706	0:16:34.231426	80.10
InceptionResNetV2	76,802	0:15:50.573017	82.43
EfficientNetV2L	125,442	0:17:19.732781	84.19
Custom CNN	272,545	0:19:43.820124	93.08

5 Conclusion

Based on gender classification from eye images, this study compares the performance of state-of-the- art deep CNN models: VGG16, ResNet152V2, InceptionResNetV2, EfficientNetV2L along with custom CNN model. The "Female and Male" dataset used to train these models. The pre-trained models shows better results for 'imagenet' for thousands of classes and millions of parameters as shown in Table 2. These models can be used to train deep neural network in short period of time with minimum number of parameters as shown in Table 3 by compromising the accuracy. It is also observed that the training time increases with the number of trainable parameters. The pre-trained models are best suited when resources are limited.

References

Majumdar, J., & Patil, B. S. (2013). A comparative analysis of image fusion methods using texture. Lecture Notes in Electrical Engineering, 221 LNEE (VOL. 1), 339–351. https://doi.org/10.1007/978-81-322-0997-3_31

Bobeldyk, D., & Ross, A. (2019). Predicting soft biometric attributes from 30 pixels: A case study in NIR ocular images. Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision Workshops, WACVW 2019, 116–124. https://doi.org/10.1109/WACVW.2019.00024

C, S. K., & Jones B, R. S. (2020). Gender Prediction from Iris Recognition using Artificial Neural Network (ANN). www.ijert.org

Singh, M., Nagpal, S., Vatsa, M., Singh, R., Noore, A., & Majumdar, A. (2018). Gender and ethnicity classification of Iris images using deep class-encoder. IEEE International Joint Conference on Biometrics, IJCB 2017, 2018-Janua, 666–673. https://doi.org/10.1109/BTAS.2017.8272755

Tapia, J., & Aravena, C. C. (2018). Gender classification from periocular NIR images using fusion of CNNs models. 2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis, ISBA 2018, 2018-January, 1–6. https://doi.org/10.1109/ISBA.2018.8311465

Thomas, V., Chawla, N. V, Bowyer, K. W., & Flynn, P. J. (2007). Learning to predict gender from iris images.

Dantcheva, A., Elia, P., & Ross, A. (2016). What else does your biometric data reveal? A survey on soft biometrics. IEEE Transactions on Information Forensics and Security, 11(3), 441–467. https://doi.org/10.1109/TIFS.2015.2480381

Abdalrady, N. A., & Aly, S. (2020, February). Fusion of multiple simple convolutional neural networks for gender classification. In 2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE), 251-256. IEEE.

Akbulut, Y., Şengür, A., & Ekici, S. (2017, September). Gender recognition from face images with deep learning. In 2017 International artificial intelligence and data processing symposium (IDAP), 1-4. IEEE.

Arora, S., & Bhatia, M. P. S. (2018, July). A robust approach for gender recognition using deep learning. In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 1-6. IEEE.

Eidinger, E., Enbar, R., & Hassner, T. (2014). Age and gender estimation of unfiltered faces. in IEEE Transactions on Information Forensics and Security, 9 (12), 2170-2179, doi: https://doi.org/10.1109/TIFS.2014.2359646.

Eyes-rtte. (23 September 2021). https://www.kaggle.com/pavelbiz/eyes-rtte

Janahiraman, T. V., & Subramaniam, P. (2019, October). Gender Classification Based on Asian Faces using Deep Learning. In 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET) (84-89). IEEE.

Keras applications. (20 October 2021). https://keras.io/api/applications/Color-feret-database. (12 October 2021). https://www.nist.gov/itl/products-and-services/Gallagher, A. C., & Chen, T. (2008, June). Clothing cosegmentation for recognizing people. In 2008 IEEE conference on computer vision and pattern recognition, 1-8. IEEE.

Gallagher, A. C., & Chen, T. (2008, June). Clothing cosegmentation for recognizing people. In 2008 IEEE conference on computer vision and pattern recognition, 1-8. IEEE.

Huang, G. B., Ramesh, M., Berg, T., & Learned-Miller, E. (2007). Labeled faces in the wild: a database for studying face recognition in unconstrained environments. Univ. Massachusetts, Amherst, MA, USA.