1st International Conference on Frontiers in Academic Research

https://www.icfarconf.com/

February 18-21, 2023, Konya, Turkey

Age Estimation from Facial Images Using Custom Convolutional Neural Network (CNN)

Gilbert George*[1], Steve Adeshina[2] and Moussa Mahamat Boukar[3]

[1]*Department of Computer Science, Nile University, Nigeria*
[2] *Department of Engineering, Nile University, Nigeria*
[3] *Department of Computer Science, Nile University, Nigeria*

*gilbertgeorge007@gmail.com *Email of the corresponding author*

Abstract – Given that aging is influenced by a variety of factors, including gender, ethnicity, environment, and others, automatic age assessment of facial images is a difficult challenge in computer vision and image analysis. Additionally, a significant amount of data and a laborious training phase are needed to estimate age from facial photos with near accuracy. In this study, we present a custom convolutional neural network-based age estimator that can almost precisely predict age from facial photos. We use the UTK facial image dataset using about 17475 images. We train the model to group the facial images into three groups which are; Child, Teenager and Adult. Compared to similar efforts, our method uses less training data while maintaining a high accuracy of 95%.

Keywords – Artificial Neural Network, Computer Vison, Convolutional Neural Network, Facial Age Classification

I. Introduction

One of the hardest tasks in the history of facial analysis is determining a person's age from a photograph. The topic of computer-based age assessment from face photographs has gained interest due to the quick development of computer vision and pattern recognition. An image of a human face can be used to acquire a variety of information, including that person's identification, emotional state, ethnicity, gender, and age. The definition of a human face image is a mixture of several facial characteristics, such as skin tone, geometric facial features, etc. All of this information is crucial when we communicate face-to-face. We are quite good at recognizing faces and face motions properly.

Numerous potential real-world applications exist for image-based automatic age estimate systems .If computers had the ability to discern the user's age, the computing environment and the sort of interaction could be tailored to the user's age also in terms of access to explicit and gambling sites, the age detection system can be used for provide access control to such resources online also by filtering the gallery database using the predicted age, age estimation from photos can be utilized to index e-photo albums as well as identify potential suspects more precisely and effectively and purchase unsuitable goods from vending machines based on their age range.

In this article, we suggest a method for teaching a CNN model to accurately deduce ages from recognized faces to classify them into three categories (Child, Teenager and Adult). The earlier efforts that are connected to our work and their shortcomings are listed in Section 2. Our operational processes and steps are detailed in Section 3. The experimental results of our work are presented in part 4, and the conclusion is provided in section 5

II. MATERIALS AND METHOD

In order to perform classification in the field of age estimation into different age groups, we have created a model based on the CNN algorithm These techniques are described next.

A. Dataset

The UTKFace dataset [2] (aligned and cropped) contains approximately 20,000 face photos with annotations for age, gender, and ethnicity. It is used in this study. We used a total of 17475 facial images.

The photographs span a wide range of variations in position, resolution, occlusion, lighting, and face expression. We selected this dataset because to its generally more uniform distributions, the variety of image properties it contains, including brightness, occlusion, and position, as well as the fact that it contains photographs of the general public. Figure 1 displays a few examples of photos from the UT Face collection. Age (in years), gender (Male-0, Female-1) and races (White-0, Black-1, Asian-2, Indian-3, and Others-4) are all tagged with a 3-element tuple for each image.

Fig.1 sample of the UTK Face Collections [2]

We focused more on the age label as the gender and ethnicity was not in the scope of this research work.

We employed the same collection of images for training and testing of our method (custom CNN).

To do this, the data sets were split into train and test groups in the ratios 70: 30. To avoid distribution mismatches during the training and testing of the models, this division was carried out while making sure that the data distribution in each division remained essentially the same. The training and test data composition in terms of age are shown in Tables 1 and 2, respectively.

B. Deep CNN

Age classification is one of the jobs that is handled utilizing the deep CNN method. A number of convolutional blocks are followed by a set of FC (fully connected) layers for classification and regression in our models' basic structures. The model receives an RGB image and scales it to 32 by 32 by 3. To address the issue of covariate shift, each architecture includes convolutional blocks, which are stacks of convolutional layers (filter size is 3x3) followed by max pooling (2x2) and dropout with value =0.3 which helps in avoiding overfitting and also to encourage independence between these deeper layers, which drops entire feature maps. The output is flattened after the convolutional blocks and then fed into the FC layers. These FC layers use a sigmoid activation function in a dense layer because the classification is non-binary and have a ReLU activation function, dropout (value between 0.2 and 0.3), batch normalization, and batch normalization. 5 folds and 50 epochs were used in the experiment's cross-validation process. The age estimate architecture is depicted in Figure 2.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_3 (Conv2D)            (None, 32, 32, 64)        832

max_pooling2d_3 (MaxPooling  (None, 16, 16, 64)        0
2D)

dropout_4 (Dropout)          (None, 16, 16, 64)        0

conv2d_4 (Conv2D)            (None, 16, 16, 32)        8224

max_pooling2d_4 (MaxPooling  (None, 8, 8, 32)          0
2D)

dropout_5 (Dropout)          (None, 8, 8, 32)          0

conv2d_5 (Conv2D)            (None, 8, 8, 32)          4128

max_pooling2d_5 (MaxPooling  (None, 4, 4, 32)          0
2D)

dropout_6 (Dropout)          (None, 4, 4, 32)          0

flatten_1 (Flatten)          (None, 512)               0

dense_2 (Dense)              (None, 256)               131328

dropout_7 (Dropout)          (None, 256)               0

dense_3 (Dense)              (None, 3)                 771

=================================================================
Total params: 145,283
Trainable params: 145,283
Non-trainable params: 0
_________________________________________________________________
```

Fig 2. model architecture

III. Results

With the use of our model on the facial images, the model was able to reach an accuracy for age classification model of 94 %

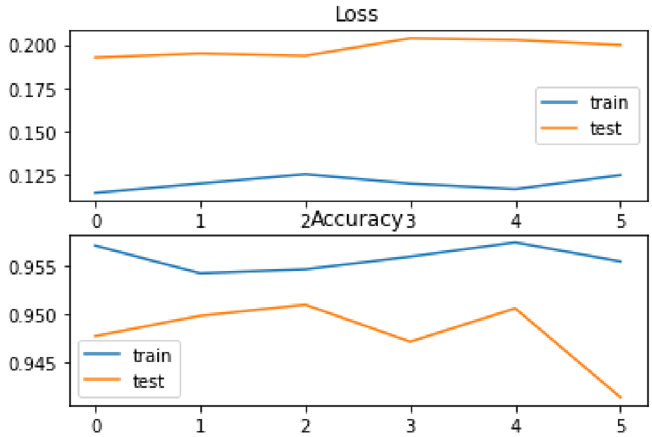

Figure 3: loss vs Accuracy epoch

A. Predictions

Below is an evaluation of our model on facial images for prediction vs actual age group.

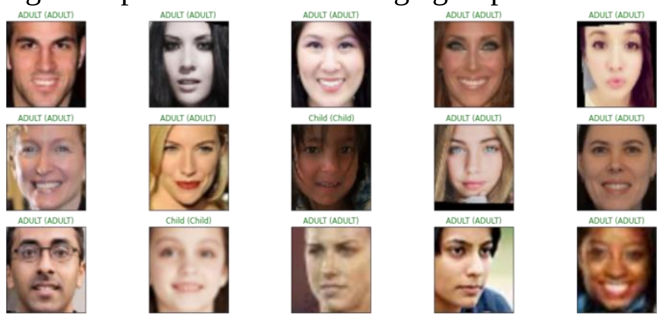

Fig 4: shows our models performance on facial images prediction vs actual.

Where predicted is in lower case and actual is in upper case and green color shows where the model predicted the facial age right and red color indicates the model failed in predicting the facial image.

B. confusion Matrix

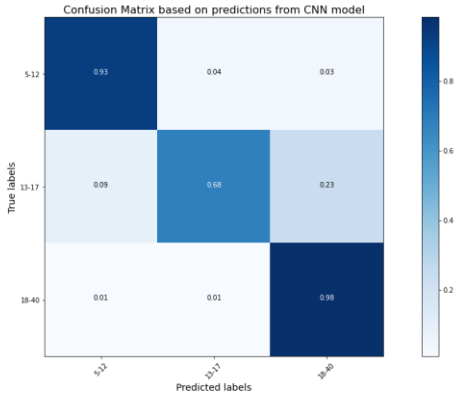

Fig 5: Confusion Matrix

From the confusion matrix above we can see the model performed well in predicting 5-12 with 93% accuracy, 13-17 with 68% accuracy and 18-40 with 99% accuracy. This tells us our model, is having a difficult time classifying teenagers as adults and children can be easily identified by humans, while with teenagers we have a bit of difficulty.

IV. Discussion

a. Comparison

Table 1: shows the past work done and their results

Work	Accuracy	Dataset
Our Model	95 %	UTK
[1]	85.5%	All-Age Face (AAF)
[2]	94.59%	FG-Net
[3]	84.8%	Aging Faces in the Wild (AGFW)

From the result gotten, it shows our work has the best accuracy from the others, and this can also be improve by ensemble our model with other pre-trained models such as Resnet-50 and inception-4.

V. Conclusion

This study put forth a classification of facial age groups with age prediction. A custom model has been implemented, for predicting age groups. We employed a custom CNN approach for age prediction.

Future work: Our model could be further enhanced, and the scope of it entails entity recognition, entity addition, and expanding the number of convolutional layers each block and fine-tuning hyper-parameters to increase the efficiency of our architecture.

.

References

[1] F. Dornaika, S. E. Bekhouche, and I. Arganda-Carreras, "Robust regression with deep CNNs for facial age estimation: An empirical study," *Expert Syst Appl*, vol. 141, Mar. 2020, doi: 10.1016/J.ESWA.2019.112942.

[2] O. Agbo-Ajala and S. Viriri, "Deep learning approach for facial age classification: a survey of the state-of-the-art," *Artificial Intelligence Review 2020 54:1*, vol. 54, no. 1, pp. 179–213, Jun. 2020, doi: 10.1007/S10462-020-09855-0.

[3] A. Singh, N. Rai, P. Sharma, P. Nagrath, and R. Jain, “Age, Gender Prediction and Emotion recognition using Convolutional Neural Network.” [Online]. Available: https://ssrn.com/abstract=3833759
[4] D. Yi, Z. Lei, and S. Z. Li, “Age estimation by multi-scale convolutional network,” Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9005, pp. 144–158, 2015, doi: 10.1007/978-3-319-16811-1_10/COVER.
[5] Z. He *et al.*, “Data-Dependent Label Distribution Learning for Age Estimation,” *IEEE Transactions on Image Processing*, vol. 26, no. 8, 2017, doi: 10.1109/TIP.2017.2655445.
[6] E. Koutsoupias and D. S. Taylor, “The CNN problem and other k -server variants,” *Theor Comput Sci*, vol. 324, no. 2-3 SPEC. ISS., 2004, doi: 10.1016/j.tcs.2004.06.002.
[7] T. J. Jun *et al.*, “TRk-CNN: Transferable Ranking-CNN for image classification of glaucoma, glaucoma suspect, and normal eyes,” *Expert Syst Appl*, vol. 182, 2021, doi: 10.1016/j.eswa.2021.115211.
[8] S. Chen, C. Zhang, and M. Dong, “Deep Age Estimation: From Classification to Ranking,” *IEEE Trans Multimedia*, vol. 20, no. 8, 2018, doi: 10.1109/TMM.2017.2786869.
[9] B. bin Gao, C. Xing, C. W. Xie, J. Wu, and X. Geng, “Deep Label Distribution Learning with Label Ambiguity,” *IEEE Transactions on Image Processing*, vol. 26, no. 6, 2017, doi: 10.1109/TIP.2017.2689998.
[10] R. Zheng, S. Zhang, L. Liu, Y. Luo, and M. Sun, “Uncertainty in Bayesian deep label distribution learning,” *Appl Soft Comput*, vol. 101, 2021, doi: 10.1016/j.asoc.2020.107046.
[11] I. Domingues, “An automatic mammogram system: from screening to diagnosis,” *Ph.D. Thesis*, no. July, 2014.
[12] J. Prajapati, A. Patel, and P. Raninga, “Facial Age Group Classification,” *IOSR Journal of Electronics and Communication Engineering*, vol. 9, no. 1, pp. 33–39, 2014, doi: 10.9790/2834-09123339.
[13] A. Bearman and C. Dong, “Human Pose Estimation and Activity Classification Using Convolutional Neural Networks,” *Stanford CS231n*, 2015.
[14] B. Zhang and Y. Bao, “Age Estimation of Faces in Videos Using Head Pose Estimation and Convolutional Neural Networks,” *Sensors 2022, Vol. 22, Page 4171*, vol. 22, no. 11, p. 4171, May 2022, doi: 10.3390/S22114171.
[15] Y. Nam and C. Lee, “Cascaded convolutional neural network architecture for speech emotion recognition in noisy conditions,” *Sensors*, vol. 21, no. 13, 2021, doi: 10.3390/s21134399.
[16] A. A. Ahmed and M. Echi, “Hawk-Eye: An AI-Powered Threat Detector for Intelligent Surveillance Cameras,” *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3074319.
[17] I. Ul Haq, A. Ullah, K. Muhammad, M. Y. Lee, and S. W. Baik, “Personalized Movie Summarization Using Deep CNN-Assisted Facial Expression Recognition,” *Complexity*, vol. 2019, 2019, doi: 10.1155/2019/3581419.
[18] O. Agbo-Ajala and S. Viriri, “Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces,” *Scientific World Journal*, vol. 2020, 2020, doi: 10.1155/2020/1289408.