Age and Gender Prediction using Deep Learning Algorithms

[1]Adnan Azam, [2]Avi Coomar, [3]Akash Bharadwaj, [4]Mahesh Kumar M R

[1, 2, 3] Student, School of CSE, REVA University, Bengaluru, India
[4]Assistant Professor, School of CSE, REVA University, Bengaluru, India
[1] adnanazam26@gmail.com, [2]r18cs506@cit.reva.edu.in , [3]akash267000@gmail.com, [4] maheshkumar.mr@reva.edu.in

Abstract

Deep Learning has gained a lot of popularity recently. One sort of Deep Learning Neural Network is the Convolution Neural Network algorithm (CNN). CNN is a strong neural network that is widely used in machine vision, image analysis and categorization, and identification, among other applications. Facial analytics from photos as variety of issues like enhanced content suggestion systems, threat monitoring using cameras, consumer specific marketing, and other disciplines. To this end, we propose an application which is capable of making predictions of age and gender on a given photograph. We train our model to the most recent Adience benchmark for predicting age and gender. The network trained on OUI-Adience collection showed some promising results, performing with the test accuracy of 95% and 80% for age and gender estimation respectively.

***Keywords* -** Image Classification, CNN, Deep Learning, Image processing, Age, Gender, Prediction, EfficientNet, Adience benchmark

1. INTRODUCTION

Facial analysis has gained a lot of attention in recent years and its study has rapidly grown by means of not only the most proficient engineers, but neuroscientists as well, because it has various possible uses in computer vision communication and autonomous access control systems. Face detection is an initial stage in digital face recognition. Face identification, on the other hand, isn't always accurate since there are several variations of photographs, such as stance variation (frontal, non-frontal), obstruction, photo alignment, lighting conditions, and facial expression [1].

1.1. Picture Processing

Picture processing is the process of improving image photos collected from cameras, airplanes, satellites, and photographs taken in daily life. Many approaches and computations are used in image processing based on analysis. Virtual pictures must be carefully developed and analysed.

Picture processing consists of two main processes that are bridged by simple steps. Photo enhancement refers to the modification of a photograph with the goal of producing higher-quality images, which may be accomplished via the use of various apps. The other approach is the most often used strategy for extracting information from a photograph. Segmentation is the process of dividing pictures into distinct components [2].

The challenge of facial recognition has been extensively researched. “Support vector machines” [17][22] Neural networks, template matching, color mapping, maximal rejection classification, and model-based detection have all been utilized. However, designing algorithms that operate for all lighting conditions, facial tones, shapes and sizes, and image backdrops are significantly more challenging. As a consequence, facial recognition remains an art as much as a science [3] [21].

The development of thoughts aids in the establishment of particular limits. Age assessment is still a multi-class challenge in which the years are categorized. People of various ages have different facial characteristics, making it challenging to combine the photographs.

Approaches for determining the age and sex of various faces are used in a variety of ways. The CNN extracts the features from the neural network. Each image gets categorized into an age group based on the provided models. The features are further analyzed and eliminated from the training modules.

2. RELATED WORK

Many researchers have suggested different deep learning techniques to classify age and gender out of an image.

A. Age / Gender Classification

In recent days, interest in the issue of extracting age-related features from face photographs, and various techniques have been proposed. A thorough review of such methodologies can be read in [4] and, more lately, in [5]. Although we are focusing on classifying the age “group” and not the exact age.

The first approaches for estimating age focused on measuring proportions between different facial features [6]. When facial features (example nose, eyes, ears, chin, mouth, etc.) have been located and their measurements of sizes and distances, between them, proportions are determined and used to classify the face into different categories of age groups using handcrafted regulations. Recently, [7] used something similar to estimate age in subjects under the age of 18. Because they necessarily require precise facial feature localization, which is one difficult problem to solve in itself as they are ill-suited for the type of images found on social media platforms.

In the first approaches for “Gender classification” [8] employed a neural network which was trained on a limited number of nearly frontal face images. They had used a machine learning algorithm which is the SVM Classifiers [9]. The Webers Local Texture Descriptor [10] was recently used for gender classification, with performance being almost perfect by the FERET bench-mark [11].

[18] used multiple convolutional neural networks for detection and alignment of facial image. Then used voting system to consolidate the predictions from various networks. Researchers in [19] used two level CNN architecture for extraction and classification out of an image. Also, they have used image processing algorithm to capture unfiltered faces before passing into the CNN model.

We use the “Adience” dataset, which has complex pictures than those provided by LFW, for age prediction., and report the performance using a better method built to utilize data from huge sample training datasets in a much better way.

3. METHODOLOGY

A. Age / Gender classification using Convolutional Neural Network

Procuring a humongous, labelled image training data set for age / gender classification from social image repositories most often presupposes either access to confidential information on the participants appearing in the images (their date of birth, gender, etc), is almost always strictly limited, or physically categorizing the photos, which is laborious and time-consuming. That's why when juxtaposed to an even larger picture categorization sets of data, dataset for predicting age / gender from true social images are rather minimal (e.g., the “ImageNet” dataset [12]). Overfitting is a major complaint when ml techniques have been used on such tiny picture resources. This challenge is exacerbated by the enormous number of model parameters in CNNs. In any of these contexts, prudence is required to prevent generalization error [13].

B. EfficientNetV2

For high-quality, speedy image classification, EfficientNets has been a benchmark. They were introduced roughly three years ago and quickly gained popularity because of the way they scaled, which provided them to learn comparatively quicker than previous networks. From the point of learning duration, EfficientNetV2 is a marked improvement over earlier versions, for a little advantage in accuracy. It adopts the paradigm of progressive learning, which entails that the imagery sizes initially minimal and gradually rise in size as the training goes ahead. The justification which is provided is that earlier versions' training capacity bogs down as picture resolution goes up.

EfficientNets employ a "depthwise convolution layer," has FLOP and lower complexity than most of the other conv layers. To solve the problem, a new layer dubbed "Fused-MB Conv layer" was proposed in a study entitled "MobileDets: Searching for Object Detection Architectures for Mobile Accelerators." In this respect, EfficientNetV2 incorporates this additional layer. Since it has a larger number of parameters, researchers can't realistically exchange all the old MB Conv layers with the fused ones.

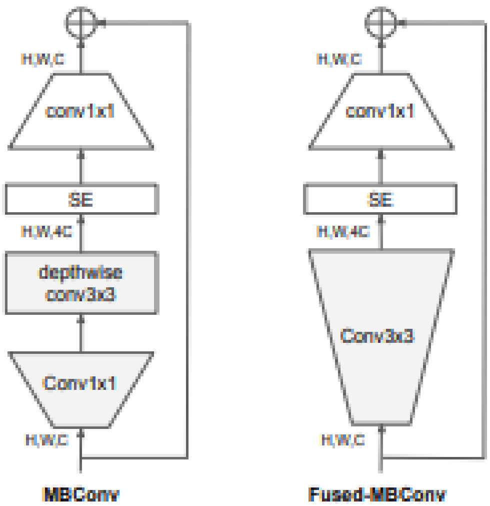

Figure 1. **Structure of MBConv and Fused-MBConv**

Using a simple compound scaling strategy, EfficientNet evenly scales up all phases. According to the authors of EfficientNetV2, this is redundant, since expanding is not intended for all of these initiatives to change performance. As a consequence, they have been using a different amplification mechanism to systematically add more layers in successive phases. They also include a sizing restriction to cap peak picture dimensions, as EfficientNets have a tendency to strongly scale up graphics, which is responsible for the characteristic concerns.

4. EXPERIMENTS

A. The Adience benchmark

Adience benchmark, tailored for age and gender categorization, is what we have used as an evaluation metrics for our CNN design [14]. Because the pictures in the Adience set were uploaded authentic without any extra filter added to it, (e.g., photographs from the LFW series [15]) or social websites (the company snapshots set [16]), the viewing contexts in these selfies are moderately free and open, indicating the humongous real-world conditions of photos consisting of face.

In 5 levels, the file comprises around 26 thousand photos of 2,284 persons. To underscore the clear edge related to the core network rather than improved preparation, these photos are presented instead of newer alignment algorithms. We applied the same network design for all the benchmark's test folds and, in reality, for both gender and age class roles. This was to make sure our outcomes are consistent and more precise, as well as to exhibit the universality of the network model provided here; after comparing, we found that it worked properly in many different situations.

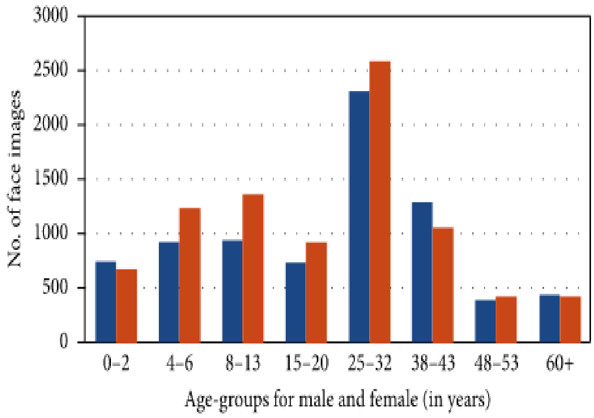

Figure 2 Number of images age wise

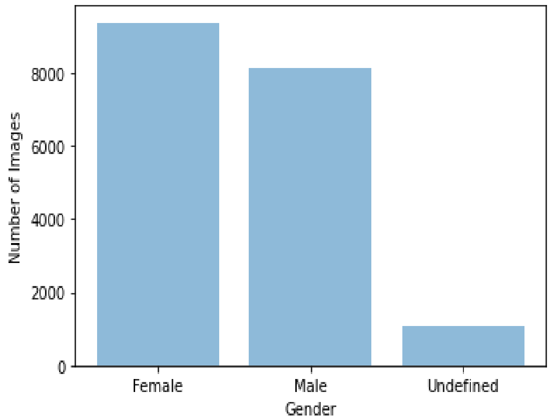

Figure 3 Gender wise number of images

5. IMPLEMENTATION

The python libraries such as Keras, TensorFlow, Matplotlib, pandas, and NumPy were utilized to build the system. Keras has certain built-in features, including activation functions, optimizers, layers, and so forth. The system backend is based on Tensorflow.

Name	Experimental tool
Hardware	Python 3 Google Compute Engine backend with Tesla P80 GPU
Software	Google Colab (Python 3.7)
Programming Language	Python
Method implementation	Keras 2.2.4 & TensorFlow 2.4.0

Table 1 Tools used for system implementation.

As we had mentioned above, the extraction of the features was done using the EfficientNet model. The tools utilized in this system implementation for experimentation are listed in the table above which shows the hardware, software, and the backend implementation. After connecting the model's last maximum pooling layer to a global average pooling, we then trained the dense layers for our dataset. On the last layer we used activation functions such as sigmoid and softmax for gender and age respectively, for converting the output layer to a vector that represents the probability distribution of a list of possibilities for 2 categories for age estimation and 8 categories for gender estimation respectively. While training, we used 13961 images in a batch size of 32, which went on for 10 epochs, an "Adam" optimizer, binary-cross-entropy along with sparse categorical cross entropy loss functions to train the CNN. Upon evaluating the performance of our custom CNN model, we then tested the efficient v2 model on 3490 test images whose results are reflected in the result section.

6. RESULT

With custom CNN sequential model using layers which are: Conv2D, MaxPoolingLayer, LayerNormalisation, Flatten Layer, and Dense Layer, accuracy obtained was:

Category	Accuracy
Gender Prediction	≈55%
Age Prediction	≈30%

Table 2.1 Accuracy with Custom CNN sequential model

With Transfer Learning using feature vectors of images with, EfficientNet V2 model, which was trained on ImageNet21k and fine-tuned on our dataset, accuracy obtained was:

Category	Accuracy
Gender Prediction (Fine-tuned)	≈95%
Age Prediction (Fine-tuned)	≈80%

Table 2.2.Accuracy with Transfer Learning

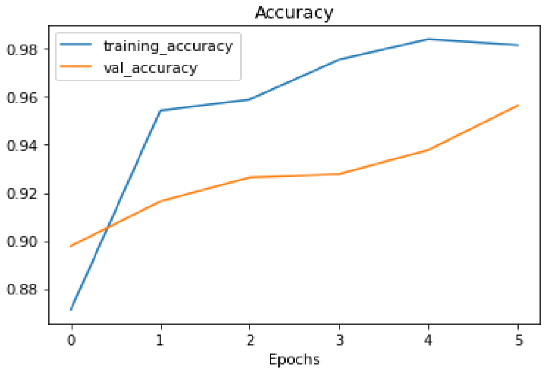

Figure 3.2. **Epoch versus Accuracy of Gender Model**

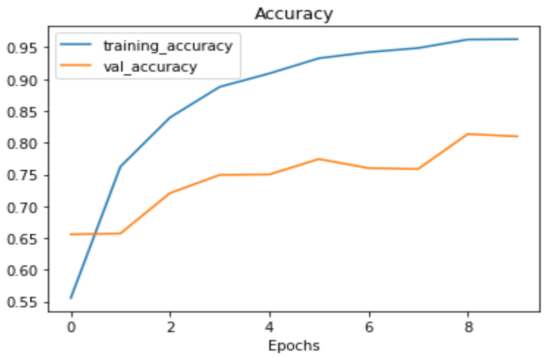

Figure 3.3 Epochs versus Accuracy of Age Model

7. CONCLUSION

The deep-CNN model we propose is derived from the base EfficientNetV2 trained on ImageNet21k, which we fine-tune with our problem-specific feature vectors on the OIU-Adience benchmark. Our dataset has enabled us to train our model on unfiltered real world images, giving it a chance to be leveraged in real world applications such as security and targeted advertisements.

We found how deep learning showed some promising results in image related problems, specially how our CNN model outdid some recent works. In the future, we aim to enhance our to provide even better accuracy and explore how it can be used in real-time applications.

8. REFERENCES

[1] Yoganand, A.V., Kavida, A.C. & Rukmanidevi. Face detection approach from video with the aid of KPCM and improved neural network classifier. *Multimed Tools Appl* 77, 31763–31785, 2018.

[2] Sonia Singla. Age and Gender Detection Using Deep Learning, 2021.

[3] Waqar Mohsin, Noman Ahmed and Chung-Tse. Face Detection Project. EE368 Digital Image Processing Spring 2002-2003.

[4] Y. Fu, G. Guo, and T. S. Huang. Age synthesis and estimation via faces: A survey. Trans. Pattern Anal. Mach. Intell., 32(11):1955–1976, 2010.

[5] H. Han, C. Otto, and A. K. Jain. Age estimation from face images: Human vs. machine performance. In Biometrics (ICB), 2013 International Conference on. IEEE, 2013.

[6] Y. H. Kwon and N. da Vitoria Lobo. Age classification from facial images. In Proc. Conf. Comput. Vision Pattern Recognition, pages 762–767. IEEE, 1994.

[7] N. Ramanathan and R. Chellappa. Modeling age progression in young faces. In Proc. Conf. Comput. Vision Pattern Recognition, volume 1, pages 387–394. IEEE, 2006.

[8] B. A. Golomb, D. T. Lawrence, and T. J. Sejnowski. Sexnet: A neural network identifies sex from human faces. In Neural Inform. Process. Syst., pages 572–579, 1990.

[9] B. Moghaddam and M.-H. Yang. Learning gender with support faces. Trans. Pattern Anal. Mach. Intell., 24(5):707– 711, 2002.

[10] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao. Wld: A robust local image descriptor. Trans. Pattern Anal. Mach. Intell., 32(9):1705–1720, 2010.

[11] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss. The feret database and

evaluation procedure for face-recognition algorithms. Image and vision computing, 16(5):295–306, 1998.

[12] Deng, Jia & Dong, Wei & Socher, Richard & Li, Li-Jia & Li, Kai & Li, Fei-Fei. (2009). ImageNet: a Large-Scale Hierarchical Image Database. IEEE Conference on Computer Vision and Pattern Recognition. 248-255. 10.1109/CVPR.2009.5206848.

[13] Gil Levi and Tal Hassner. Age and Gender Classification using Convolutional Neural Networks.

[14] E. Eidinger, R. Enbar, and T. Hassner. Age and gender estimation of unfiltered faces. Trans. on Inform. Forensics and Security, 9(12), 2014.

[15] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical report, Technical Report 07-49, University of Massachusetts, Amherst, 2007.

[16] A. C. Gallagher and T. Chen. Understanding images of groups of people. In Proc. Conf. Comput. Vision Pattern Recognition, pages 256–263. IEEE, 2009.

[17] Cortes, C., Vapnik, V. Support-vector networks. *Mach Learn* **20,** 273–297 (1995).

[18] Khaled Rahman Hassan , Israa Hadi Ali, Age and Gender Classification using Multiple Convolutional Neural Network, 2nd International Scientific Conference of Al-Ayen University (ISCAU-2020), 2020 IOP Conf. Ser.: Mater. Sci. Eng.

[19] Olatunbosun Agbo-Ajala, Serestina Viriri, "Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces", *The Scientific World Journal*, vol. 2020.

[20] T. Shreekumar, N. V. Sunitha, N. Suhasini, K. Suma, K. KArunakara, “Blur and Noise Removal from the Degraded Face Images for Identifying the Faces Using Deep Learning Networks”, International Conference on Artificial Intelligence and Sustainable Engineering 2022, pp 325–341, DOI: 10.1007/978-981-16-8546-0_27

[21] Periasamy, Keerthika, et al. "A proactive model to predict osteoporosis: An artificial immune system approach." *Expert Systems* 39.4 (2022): e12708.

[22] Basha, Syed Muzamil, and Dharmendra Singh Rajput. "Evaluating the impact of feature selection on overall performance of sentiment analysis." *Proceedings of the 2017 International Conference on Information Technology*. 2017.