International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST21835

Age and Gender Detection Using CNN

Prof. Jaydeep Patil[1], Rohit Thombare[2],Yash deo[2], Rohit Kharche[2], Nikhil Tagad[2]

[1]Assistant Professor at Information Technology Department, AISSMS Institute of Information technology, Pune, Maharashtra, India

[2]B.E., Information Technology Department, AISSMS Institute of Information Technology, Pune, Maharashtra, India

Article Info
Volume 8, Issue 3
Page Number : 29-33

Publication Issue
May-June-2021

Article History
Accepted : 01 May 2021
Published : 05 May 2021

ABSTRACT

In recent years, much effort has been put forth to balance age and sexuality. It has been reported that the age can be accurately measured under controlled areas such as front faces, no speech, and stationary lighting conditions. However, it is not intended to achieve the same level of accuracy in the real world environment due to the wide variation in camera use, positioning, and lighting conditions. In this paper, we use a recently proposed mechanism to study equipment called covariate shift adaptation to reduce the change in lighting conditions between the laboratory and the working environment. By examining actual age estimates, we demonstrate the usefulness of our proposed approach.

Keywords : Face Detection, Skin Colour Segmentation, Face Features extraction, Feature's recognition, Fuzzy rules.

I. INTRODUCTION

Normally, in a society, the process of using a guard for manual inspection may not offer the best security for either your facility or for your staff. Traditional checking is comparatively tough due to the steps involved, as for each vehicle the guard must carry out some inspections, including ID card check for all owners of the vehicle and checks of its contents. For good security to be maintained, and in busy periods, this may demand the guard having to repeat the process several times per day and never failing in the checks. Adding electronic support such as card readers can verify the card to be authentic, but it still needs the guard to visually validate the person, and that all adds more time to the process.

Age and gender play fundamental roles in social interactions. Language's reserve different salutations and grammar rules for men or women, and very often different vo-vocabularies are used when addressing elders compared to young people. Despite the basic roles these attributes play in our day-to-day lives, the ability to automatically estimate them accurately and reliably from face images is still far from meeting the needs of commercial applications. This is particularly perplexing when considering recent claims to super-human capabilities in the related task of face recognition.

Face detection is a computer technology being used in a variety of applications that identifies human faces in digital images. Face-detection algorithms focus on the detection of frontal human faces. It is a type of application classified under "computer vision" technology. It is the process in which algorithms are developed and trained to properly locate faces or objects, in images. These can be in real-time from a video camera or from photographs. It is based on the "Haar" Wavelet technique to analyse pixels in the image into squares by function. This uses machine learning techniques to get a high degree of accuracy from what is called "training data".

This uses "integral image" concepts to compute the "features" detected. "Haar" Cascades uses the "Adaboost" learning algorithm which selects a small number of important features from a large set to give an efficient result of classifiers.

In this paper, we attempt to close the gap between automatic face recognition capabilities and those of age and gender estimation methods. To this end, we follow the successful example laid down by recent face recognition systems: Face recognition techniques described in the last few years have shown that tremendous progress can be made by the use of deep convolutional neural networks (CNN). We demonstrate similar gains with a simple network architecture, designed by considering the rather limited availability of accurate age and gender labels in existing face data sets.

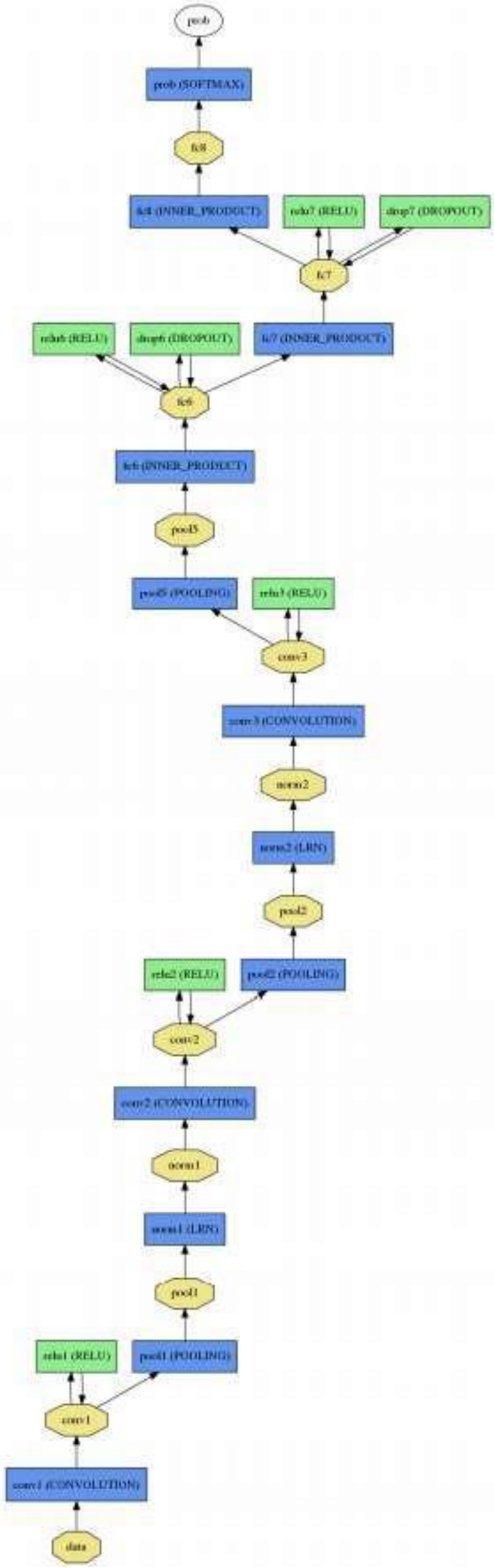

Figure 1. Full schematic diagram of our network architecture.

II. METHODS AND MATERIAL

1. A CNN for age and gender estimation:

Gathering a large, labelled image training set for age and gender estimation from social image repositories requires either access to personal information on the subjects appearing in the images (their birth date and gender), which is often private, or is tedious and time-consuming to manually label. Data-sets for age and gender estimation from real-world social images are therefore relatively limited in size and presently no match in size with the much larger image classification data-sets (e.g., the ImageNet dataset) . Overfitting is a common problem when machine learning-based methods are used on such small image collections. This problem is exacerbated when considering deep convolutional neural networks due to their huge numbers of model parameters. Care must therefore be taken in order to avoid overfitting under such circumstances.

Network architecture

Our proposed network architecture is used throughout our experiments for both age and gender classification. It is illustrated in Figure 2. A more detailed, schematic diagram of the entire network design is additionally provided in Figure 1. The network comprises only three convolutional layers and two fully-connected layers with a small number of neurons. This, by comparison to the much larger architectures, applied. Our choice of a smaller network design is motivated both by our desire to reduce the risk of overfitting as well as the nature of Figure I. Full schematic diagram of our network architecture. Please see the text for more details. of the problems we are attempting to solve: age classification on the Audience set requires distinguishing between eight classes; gender only two. This, compared to, e.g., the ten thousand identity classes used to train the network used for face recognition.

All three-color channels are processed directly by the network. Images are first rescaled to 256 x 256 and a crop of 227 x 227 is fed to the network. The three subsequent convolutional layers are then defined as follows.

1. 96 filters of size 3x7x7 pixels are applied to the input in the first convolutional layer, followed by a rectified linear operator (ReLU), a max-pooling layer taking the maximal value of 3 x 3 regions with two-pixel strides, and a local response normalization layer.
2. The 96 x 28 x 28 output of the previous layer is then processed by the second convolutional layer, containing 256 filters of size 96 x 5 x 5 pixels. Again, this is followed by ReLU, a max-pooling layer, and a local response normalization layer with the same hyperparameters as before.
3. Finally, the third and last convolutional layer operates on the 256 x 14 x 14 blobs by applying a set of 384 filters of size 256 x 3 x 3 pixels, followed by ReLU and a max-pooling layer. The following fully connected layers are then defined by:
4. A first fully connected layer that receives the output of the third convolutional layer and contains 512 neurons, followed by a ReLU and a dropout layer.
5. A second fully connected layer that receives the 512- dimensional output of the first fully connected layer and again contains 512 neurons, followed by a ReLU and a dropout layer.
6. A third, fully connected layer maps to the final classes for age or gender.

Finally, the output of the last fully connected layer is fed to a soft-max layer that assigns a probability for each class. The prediction itself is made by taking the class with the maximal probability for the given test image.

III. EQUATIONS

1. SSR-Net expects the input to be tensor size: N x 64 x 64 x 3, where N is face number, 64x64 height and width equally and 3 represents RGB. Individual values per tensor should be measured at [O ... 1]. Please note that the call function cv.normalize (blob [i,:,:,:], None, alpha= 0, beta= 255, norm_type = cv.NORM_MINMAX) performing the required customization.
2. Gil Levi and Tal Hassner's ConvNet expects the input to be tensor size: N x 3 x 227 x 227, where N has a face value, 3 means the RGB channels and 227x227 are the same height and width. Each of the channels in the tensor should say 0 but should not be rated. Please note the limitations caller I .0 and mean (78.4263377603, 87.7689143744, J 14.895847746) on the work phone cv.dnn.blobFromlmages.

IV. RESULTS AND DISCUSSION

While implementing this project we analysed different articles and models to estimate human gender and age by image.

We have discovered that there are a lot of good models with high accuracy that are yet too big and slow to compute.

On the other hand, there are some small models with lower accuracy that could be used for real-time video processing.

We have successfully used two such models for real-time estimation of age and gender using only average CPU:

- SSR-Net by Tsun-Yi Yang, Yi-Hsuan Huang, Yen-Yu Lin, Pi-Cheng Hsiu, Yung-Yu Chuang.
- ConvNet by Gil Levi and Tai Hassner.

TABLE I

Sr. No	Survey for this Paper		
	Title	**Authors**	**Description**
	Face Recognition using Haar Cascade Classifier	Varun Garg AndKritika Garg	Face detection with high efficiency using OpenCV library'sHaar cascade classifier.
2	Real time Face detection and tracking using Haar classifiero n SoC	RN Daschoudhr y Rajshr ee Tripat hi	Real time face detection from high definition Video through raspberry pi
3	Gender and	IEEE 2014	The research on
	Age recognition for video	Conference	CNN and ML ensures data extraction and
	analyti cs Solutio ns		complete mappingof nodes, by identifying,
	--- IEEE 2014 paper		analysing and interpreting the suitable Image of
			human
4	Crowd counting with respectto age and gender by using faster CNN detection		It makes use of "Model Replication" framework, in which two sequences are matched at mapping image pixel level and thematching results are grouped for a final decision. It further makes use of ranking of those groups based on pattern.

The easy availability of huge image collections provides modem machine learning-based systems with effectively endless training data, though this data is not always suitably labelled for supervised learning. Taking an example from the related problem of face recognition we explore how well deep CNN performs on these tasks using Internet data. We provide results with a lean deep-learning architecture designed to avoid overfitting due to the limitation of limited labelled data. Our network is "shallow" compared to some of the recent network architectures, thereby reducing the number of its parameters and the chance for overfitting. We further inflate the size of the training data by artificially adding cropped versions of the images in our training set. The resulting system was tested on the Audience benchmark of unfiltered images and shown to significantly outperform the recent state of the art. Two important conclusions can be made from our results.

First, CNN can be used to provide improved age and gender classification results, even considering the much smaller size of contemporary unconstrained image sets labelled for age and gender. Second, the simplicity of our model implies that more elaborate systems using more training data may well be capable of substantially improving results beyond those reported here

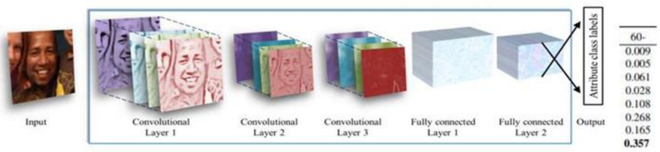

Table 1 : Age estimation confusion matrix on Audience benchmark

V. CONCLUSION

Though many previous methods have addressed the problems of age and gender classification, until recently, much of this work has focused on constrained images taken in lab settings. Such settings do not adequately reflect appearance variations common to the real-world images on social websites and online repositories. Internet images, however, are not simply more challenging: they are also abundant.

VI. REFERENCES

[1]. A. A. Zaidan, B. B. Zaidan , A. Al-Haiqi, M. L. M. Kiah, M. Hussain, and M. Abdulnabi , "Evaluation and selection of open-source EMR software packages based on integrated AHP and TOPSIS," J. Biomed. Inform., vol. 53,pp. 390--404,2015.

[2]. T. Ahon en, A. Hadid, and M. Pietikainen , "Face description with local binary patterns: Application to face recognition," IEEE Trans. Pattern Anal. Mach. Intel!., no. 12, pp. 2037- 2041, 2006.

[3]. R. Sharma, T. S. Ashwin , and R. M. R. Guddeti, " A Novel Real-Time Face Detection System Using Modified Affine Transformation and Haar Cascades," in Recent Findings in Intelligent Computing Techniques, Springer, 2019, pp. 193- 204.

[4]. M. Hussain, A. Al-Haiqi, A. A. Zaidan, B. B. Zaidan, M. L. M. Kiah, N. B. Anuar, and M. Abdulnabi, "The landscape of research on smartphone medical apps: Coherent taxonomy, motivations, open challenges and recommendations ," Comput. Methods Programs Biomed., vol. 122 , no. 3, pp. 393-408 , 2015.

[5]. N. Kalid, A. A. Zaidan, B. B. Zaid an, 0 . H. Salman, M. Hashim, and H. Muzammil, "Based real time remote health monitoring systems: A review on patients prioritization and related" big data" using body sensors information and communication technology ," J. Med. Syst., vol. 42

Cite this article as :

Prof. Jaydeep Patil, Rohit Thombare, Yash deo, Rohit Kharche, Nikhil Tagad, "Age and Gender Detection Using CNN", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 8 Issue 3, pp. 29-33, May-June 2021. Available at

doi : https://doi.org/10.32628/IJSRST21835

Journal URL : https://ijsrst.com/IJSRST21835