

Face age and gender recognition based on improved VGGNet algorithm

Yulin Li

*School of Mathematics and Statistics, Nanjing University of Information Science & Technology,
Nanjing, 210044, China*

(Received February 20 2019, accepted June 24 2019)

Abstract. Recognition of age and gender based on face image is one of the hotspots of current artificial intelligence research. In this paper, an improved VGG+SENet algorithm is proposed to simplify the identification of age and gender algorithm by simplifying VGGNet model, improving the loss function and embedding the SENet module. Compared with other models, the improved network structure and loss function model proposed in this paper can quickly and accurately obtain output recognition results. Experimental results on multiple benchmark face datasets show that the proposed improved VGG+SENet algorithm has higher recognition accuracy than other related models based on deep learning.

Keywords: VGGNet, SENet, Age estimate, Gender identification

1. Introduction

Computer vision technology is widely used in various fields. Among them, face recognition [1] has made an important breakthrough in computer vision technology. At present, face recognition has been widely used in real life, such as identity authentication, network payment, public security monitoring, image tracking and so on. Our research interests focus on apparent age and gender estimates. Age and gender classification play a very important role in social life. According to the classification, we can know whether the person being contacted is young or old, whether it is “Ms.” or “Mr.”. Facial expressions have an important influence on the ability to estimate these individual classifications. The current facial expression analysis and recognition model capabilities are still far from meeting the needs of commercial applications [2].

VGGNet is a well-known convolutional neural network model applied to image classification. In 2018, He et al. proposed a facial expression recognition method based on VGGNet deep convolutional neural network for the low recognition rate of traditional convolutional neural networks in facial expression database. With a deeper network structure and 3*3 small convolution kernels and 2*2 small pool cores, the recognition rate is significantly improved, and the number of parameters is only slightly larger than the shallow layer. Shirkey et al.[4] combine feature-based gender identification methods with histogram-based age prediction methods to achieve desired goals. Through actual age estimation experiments, they demonstrate the effectiveness of the proposed method. In 2019, Chen et al.[5]proposes a face recognition algorithm based on SVM combined with VGG network model extracting facial features, which can not only accurately extract face features, but also reduce feature dimensions and avoid irrelevant features to participate in the calculation. Ghazi et al.[6] presented a comprehensive study to evaluate the performance of deep learning based face representation under several conditions including the varying head pose angles, upper and lower face occlusion, changing illumination of different strengths, and misalignment due to erroneous facial feature localization. Rothe et al.[7] used deep learning on the IMDB-WIKI dataset to solve the estimation of facial age in static face images. With the development of computer hardware and artificial intelligence technology, even under the conditions of large-scale training data, the computing power can satisfy the face recognition algorithm technology of practical application.

Aiming at the complexity of the existing deep learning neural network, the VGGNet network structure of the convolutional neural network is improved, the network structure is simplified, the loss function is improved, the SENet module is embedded, and a face age and gender recognition based on the improved VGG+SENet algorithm is proposed. Experiments show that under the large-scale data set, the improved VGG+SENet deep learning algorithm has a high recognition rate and can meet the requirements of actual face recognition applications.

2. VGGNet neural network introduction

VGGNet is a deep convolutional network jointly developed by Oxford Visual Geometry Group and DeepMind. In 2014, it won the second place in the classification project and the first place in the positioning project in the ILSVRC(ImageNet Large Scale Visual Recognition Challenge) competition, and the image feature extraction ability is very good, and its performance in multiple migration learning tasks is excellent for GoogLeNet. Therefore, extracting features from images, the VGG model is the preferred algorithm. The disadvantage of the VGG model is that there are many parameters, and the parameter quantity is more than 138 M. It requires a large storage space and the training time is relatively long. The model structure of VGGNet [8] is shown in Figure 1.

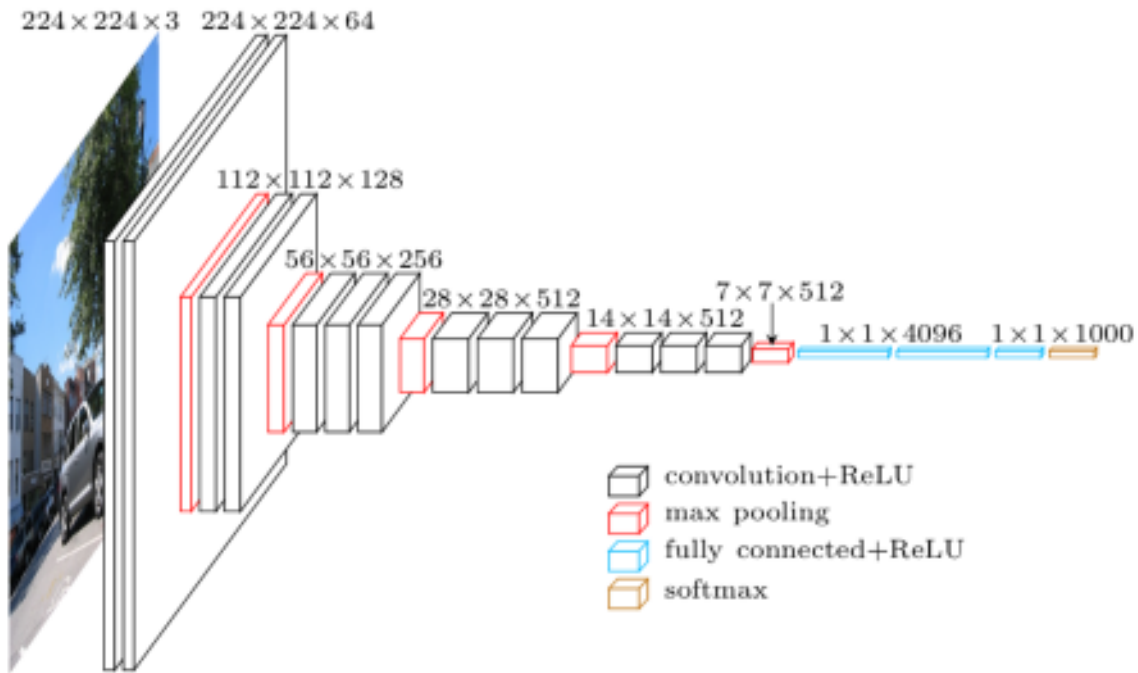


Fig. 1.VGGNet model structure

The structure of the VGGNet model is described as follows:

- 1) The input to the network is 224×224 RGB pictures, all of which are averaged.
- 2) The network has a total of 5 maximum pooling layers and 13 convolution layers. Three fully connected layers and one SoftMax classifier layer.
- 3) The convolution kernel has a size of 3×3 in the convolutional layer, a step size of 1 (stride=1), and a complement of 0 (pad=1).
- 4) The pooling layer uses MaxPooling, but not all convolution layers have a pooling layer. The pooling window is 2×2 and the step size is 2, that is, non-overlapping pooling is adopted.
- 5) All hidden layers are equipped with a ReLU layer.
- 6) After the first and second fully connected layers, dropout technology is also used to prevent network overfitting.

3. Face recognition based on improved VGG+SENet algorithm

3.1 Insufficient VGGNet network

Because VGGNet has good portability and promotion characteristics, VGGNet is chosen as the neural network model for face recognition. VGGNet has excellent image extraction effects, but has the following shortcomings:

- 1) There are many network layers, and the amount of calculation is large during training, and the convergence is slow.