Convolutional neural networks

Fangfang Li

May 25, 2018

1 Introduction

Deepgaze uses the convolution neural network (CNN) to estimate the head posture and gaze direction, and carries out skin detection, motion detection and tracking through reverse projection.

2 Neural Networks

Convolutional neural networks (CNN or ConvNet) are feedforward artificial neural networks, as shown in Figure 1. The connection pattern of neurons in the network is inspired by the animal's visual cortex structure, in which independent neurons respond to mutually covering neurons, and they together cover the entire visual area. Convolutional networks are inspired by physiological processes and are multi-layered perceptual variables designed to use only a minimal amount of preprocessing [1]. It has a wide range of applications in image and video recognition, recommendation systems, and natural language processing. Table 1 is the structure of the proposed local network. We used average pooling as a method of the position- sensitive ROI pooling. The output value of b-th bin after the pooling is calculated as:

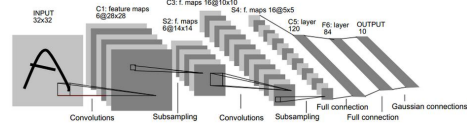$$r(b) = \frac{1}{E} \sum_{(x,y) \in bin(b)} C_b(x_0 + x, y_0 + y) \qquad (1)$$

Figure 1: Convolutional neural networks.

Table 1: The structure of the local refine network.

Layer name	Output size	Layer size
conv6	P × P	1 × 1, 2048, 1
conv7	P × P	1 × 1, 2048, 1
PS-ROI pooling	B× B × N	

References

[1] S. Lawrence, C. L. Giles, Ah Chung Tsoi, and A. D. Back. Face recognition: a convolutional neural-network approach. *IEEE Transactions on Neural Networks*, 8(1):98–113, 1997.