Lecture 2

Fourth stage

Medical Imaging Processing II

Computer vision and image analysis

By

Asst. Prof. Dr. Mehdi Ebady Manaa
Asst. Lect. Lubna Ali

Computer vision and Image Analysis

It is one of the most expanding fields of computer science in recent years, which aims to make computers capable of understanding and analyzing digital images and videos in the same way that humans vision system does. Until recently, the computer vision did not possess high capabilities, but after the tremendous leap in many fields, the skills of the computer vision grew and outperformed the capabilities of human vision in some times. Computer vision has multiple tasks such as image analysis, image processing, and extraction of high-dimensions data from images. Therefore, computer vision can be used in many applications such as image classification, object identification, age estimation, analysis of the effects of aging on the human face, and face verification and so on.

So, the machine learning approach seeks to build algorithms that gain experience through training as the human brain does. On the computer vision level, machine learning algorithms work on pattern recognition. So, if we want to train a computer to understand visual data, we need to feed the computer with thousands or millions of pre-categorized images. Then you need algorithms or relationships through which the computer searches for patterns common to the labelled images. This process is called “supervised learning”, which uses training examples to obtain relationships between patterns and outcomes. For examples

1. Face detection

Face detection is the process of determining the position of a face within an image or within a video clip and generating a rectangular boundary box around it regardless of its expression. Face detection is the pre-process that considers a building block of many computer vision applications that rely on the human face. Generally, there are numerous techniques used to perform this task, one of them uses Histograms of Oriented Gradient (HOG) descriptors that supply outstanding performance comparing to other available feature sets. Generally, there are four main steps followed to achieve HOG features which are as follows:

- ***Step 1:*** Compute the Gradient Images by apply oriented edge filter [-1, 0, 1] on the horizontal and vertical input RGB colour space image. Figure 1 shows the oriented edge filter.

$$D_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$$

$$D_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Figure 1: The horizontal and vertical oriented edge filters(the Dx is the horizontal kernel, and Dy is the vertical kernel).

When applying the oriented edge filters on the image, the result getting as in Figure 2.

Figure 2: Applying the oriented edge filter on the face image. (a) original image (b)image with applying horizontal kernel, (c) image with the vertical kernel, (d) using a horizontal and vertical kernel

- ***The second step*** in extracting the HOG features is the orientation binding which is based on the number of gradient computation obtained values. Where the face image dividing into cells of pixels, each pixel cell fling histogram channel weighted vote which is orientation based. The linear gradient vote is 9 bins of orientation in the range 0^o-180^o. Figure (3) shows the principle of orientation binning.

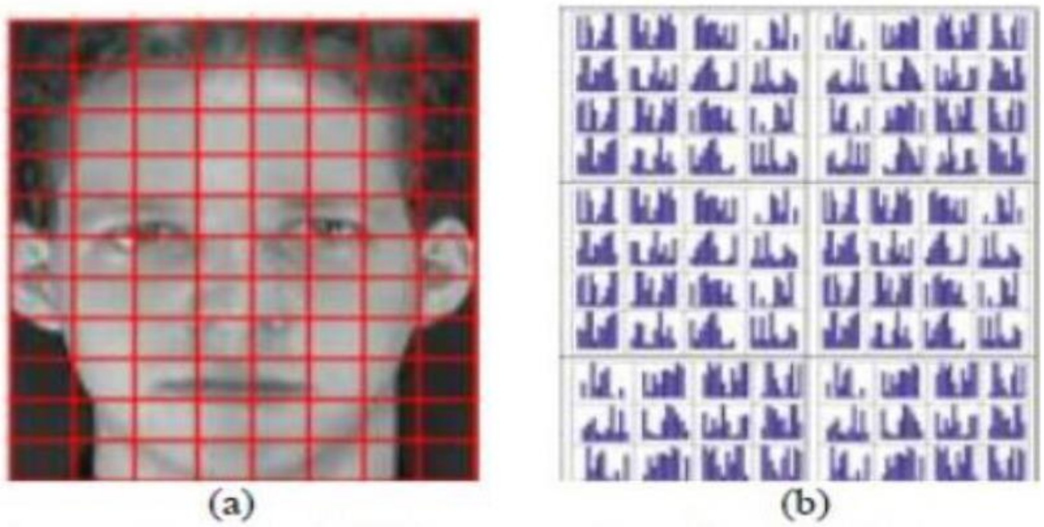

Figure 3. Orientation binning operation, (a) partition face image into cells, (b) compute cell histograms

- ***The third step is*** "Descriptor blocks" , in which the adjacent cells grouped into a larger spatial space called a block. The normalization applied for all cells within the block to reduce the effect of different contrast and illumination. It is important to mention the kinds of blocks available, there are two kinds which are rectangular R-HOG and circular C-HOG blocks. R-HOG blocks consist of a square or rectangular grids of cells. The rectangular block characterized by several cells forming each block, the number of pixels forming each cell, and each histogram bins or channels. The most appropriate parameters for the human face were noticed are 16x16 block pixels dividing into four 8x8 cells of pixels and 9 histogram bins.

- ***The fourth step*** is parameters normalization or block normalization.

Figure 4 shows HOG features extracted from the face image.

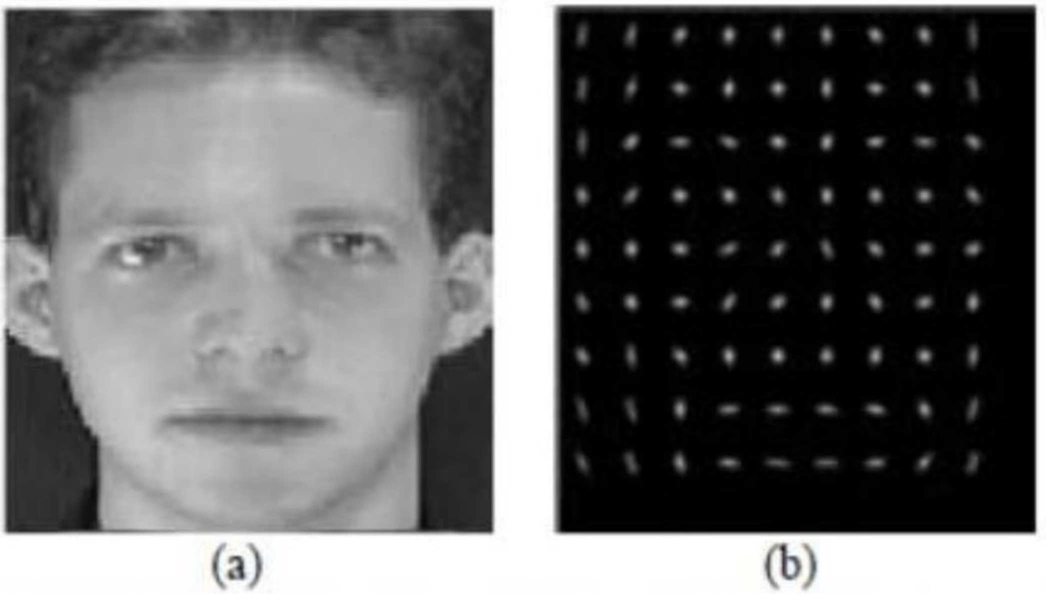

Figure 4 face image against its HOG features, (a) sample face image. (b) HOG features of the face image

2. Facial landmarks

In computer vision, facial landmark detection is the process that tries to determine the prominent areas along the face, such as position of the jaw , right and left eyebrow , right and left eye , nose, and mouth . Facial landmark detection is an important and essential stage for facial analysis methods and many applications such as face swapping, facial expression recognition, face recognition, face alignment, emotion recognition, and others. Figure 5 shows of images after detecting facial landmarks.

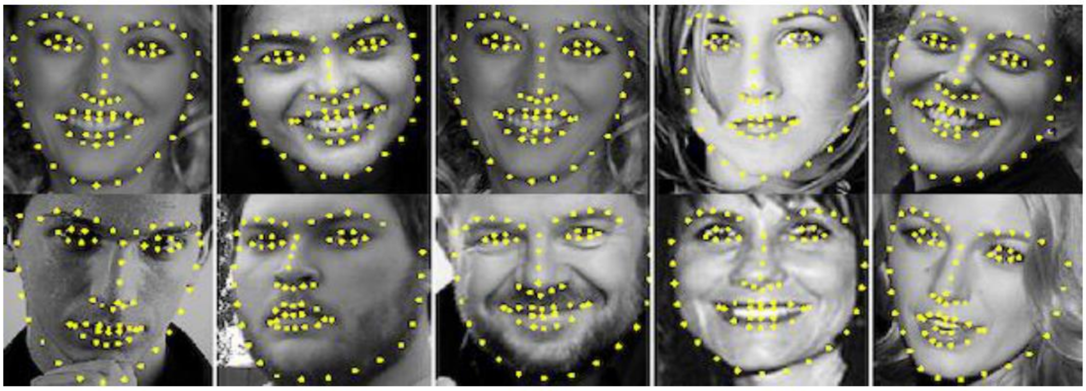

Figure 5: Some of the images after detecting facial landmark