Abstract

Image content analysis is a very challenging and complex topic, covering a broad spectrum of various tasks. In the case of image retrieval, an important aspect is to extract characteristic features describing the image or objects contained in it and express them in the form of numerical descriptors. Such descriptors should represent the crucial elements of the analyzed images and provide an effective way of searching for similar images in the database. This work aims to investigate the possibilities of deep neural networks in terms of developing a descriptor generation method for Content-Based Image Retrieval applications. The essential point is to determine the most representative areas of images. In this thesis, we present a new hybrid approach to achieving this aim, proposing descriptors based on neural activations of convolutional neural networks. Activations from both convolutional and fully-connected layers are taken into account. In the case of convolutional layers, only these activations that have the most influence on the decision-making process are included. In a series of numerical experiments conducted on the commonly known IMAGENET1M dataset, it was demonstrated that the proposed descriptors help in retrieving images which are similar not only semantically but also with respect to the secondary image characteristics, like background, textures, color distributions, etc. The proposed descriptors were also compared with other solutions that can be found in the literature, as well as with descriptors constructed solely with fully-connected or convolutional layer activations. In a majority of cases, the proposed method provided the best results, according to various measures proposed in the thesis. These measures are: the L_1 distance between descriptors, the L_1 distance between the Gramm matrices, and the overlap area between Parzen estimators of color distributions. Additionally, research concerning the interpretability of deep neural networks was carried out. It turned out that the proposed descriptors may be clustered into groups that match the hierarchy of classes provided in the considered dataset. This fact may be helpful in enhancing the classification process in the deep neural network and is a promising future research direction on this subject.