Neurocomputing 000 (2018) 1–9

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Object detection and recognition via clustered features

Marcin Woźniak*, Dawid Połap

Institute of Mathematics, Silesian University of Technology, Kaszubska 23, Gliwice 44-100, Poland

ARTICLE INFO

Article history:
Received 29 June 2017
Revised 9 July 2018
Accepted 11 September 2018
Available online xxx

Communicated by Liang Lin

MSC:
92B20
68T05
68T10
68U10

Keywords:
Convolutional Neural Networks
Adaptive systems
Automated Decision Support
Image processing

ABSTRACT

The analysis of 2D images consists of two processes: detection and recognition of detected objects. Both stages allow for numerous applications in practical purposes, including detection of small objects and people with their appearance. The methods we can implement for these can benefit from fusion of approaches. In this article, we propose detection method based on analysis of the number of clusters of points in conjunction with Convolutional Neural Network as a final classifier. Proposed method of determining the clusters of points is based on a combination of modeled graphics processing with fuzzy logic. The proposed architecture of detection and classification has been tested and compared to other approaches in this field to show the efficiency and draw conclusions for further development.

1. Introduction

Typical vision algorithms are implemented in many different areas in today's world. We can find application for tee identification by use of packet entropy and fuzzy SVM as proposed by Wang et al. [33]. There are approaches developed for self supervision as proposed by Wang et al. [32] and systems which use discriminant graph models like [18]. In general the entire process of operation of this type of methods is comprised of several steps. At the beginning, the image should be modified in such a way that it can be easier for machine to analyze it. For that purpose, various image processing techniques such as blurring, sharpening or edge detection are applied. On the basis of modified image a plurality of image features are extracted and analyzed for possible shapes. This process serves as detection of objects. In the next step, the shape and found features are subjected for classification. Final stage of the contemporary technique of most vision algorithms is decision of the classifier on recognized objects in the image.

Even though the methods of computer vision are composed of several different phases, in practice these can be separated. There are examples of applications where only detection phase is necessary. Motion detectors are one of these, where rapid processing of video frames is necessary for video supervision, i.e. for progressive face recognition [17] or facial emotion recognition [39]. A similar situation exists with methods of object recognition. In factories, each product can be automatically verified using images from production line. In case of fruits, a complete computer vision algorithm can be implemented for recognition, but also detection of any defects in peel, shape, etc. Similar example is an airport, where sensors capture human faces, and then they are compared using information from a database of prohibited passengers.

1.1. Related works

Automated detection of objects over graphic processing is desired in the industry for decades. Already in the 70s of the twentieth century, scientists began to work on detection of simple shapes including oval and round in order to detect people's faces. In [35] was showed that edge detection is an important element of analysis process. In his research he showed that shape location acting on an analysis of adjacent points allows scaling the edges to eliminate their quantity. In the first study of detection, an important driver of activity was an analysis of video file. Simple method for object detection and calculation of their distances by analyzing frames of video files was presented in [10], where proposed solution was devoted to avoiding obstacles in exploring vehicles used in the space. Fast development of different methods of graphic

* Corresponding author.
E-mail addresses: Marcin.Wozniak@polsl.pl (M. Woźniak), Dawid.Polap@polsl.pl (D. Połap).

https://doi.org/10.1016/j.neucom.2018.09.003
0925-2312/

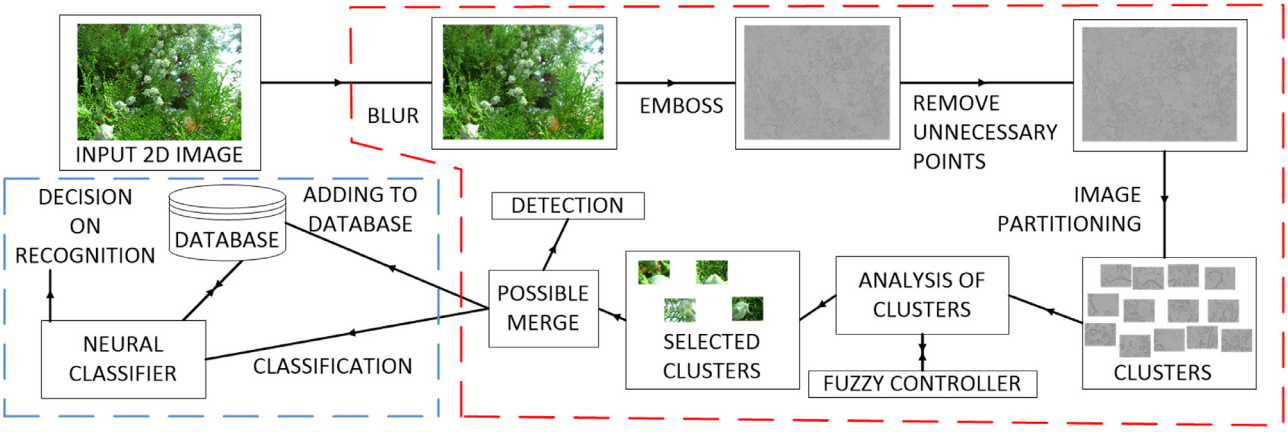

Fig. 1. Graphical representation of a model of the proposed method. Stages of detection are marked using red box and object recognition using blue one.

processing allowed to implement these techniques. For example, in [30] was presented an analysis of gray level co-occurrence matrices as an interesting solution for combining information possible to obtain by a variety of graphic filters such as gray level histograms, edge detection and texture analysis. The following years brought new aspects to the analysis of characteristics of 2D images. The flourishing of finding features in images brought primarily some algorithms such as SIFT [20] and SURF [1]. Moreover, the image does not always has been analyzed as a whole but very often as divided into smaller parts using pyramidal representation as proposed in [5]. The pyramid representation means a combination of blurring and reducing the image for the purpose of noticing changes.

Shape recognition was being developed almost parallel to detection. At the beginning of the 80s of the last century, in [2] was proposed a procedure of matching simple geometric shapes (triangles or trapezoids) using graph theory. Again, in [9] was discussed an approach by using moment invariants in recognition of affine-deformed objects. Their research showed that the use of the second and the third–order momentum is sufficient for these requirements. Furthermore, in [11] was presented the idea of using generalized Hough transform. The influence of external factors such as noise and obstruction to recognition has been described and analyzed due to the occurrence of misidentification. The late 90s marked the beginning of the common use of neural networks as classifiers for all kinds of graphical objects. This approach caused that data extraction from images had to be done in a certain way – quality of the data should be the most accurate, but their number as small as possible. This, together with long process of training, have become the most important problems in graphic recognition by the use of neural networks. In [24] neural network is used in two ways in one model of face recognition. The input image was analyzed by the first network that verifies possible rotation of the head and returns a file which is converted and placed on the entry of the second network for further identification. Extraction can be performed using the principal component analysis, which is based on trading in the coordinate system in order to maximize variance of the data. This solution of data extraction for the neural network was presented in [7]. In addition, authors suggested to use techniques of reducing the data size to fit the number of inputs to the network. One of the latest solutions in object recognition was presented in [36], where the idea to identify type of activity using video stream was based on devoted clustering (Fig. 1).

Recent advances in deep learning techniques can help on assisted image processing to adjust the image to extract particular features as presented in [37]. Similarly deep learning technique can be implemented in a framework for synthesis of objects motions as presented in [13]. Excellent examples of deep learning techniques devoted for image processing are Convolutional Neural Networks. In [26] was proposed a simplified convolutional architecture composed for rough sketches processing, while in [16] was proposed an application of Convolutional Neural Networks used to purify shapes of objects to be embedded in images.

In this article we present a novel two stage model for detection of important features from images. The proposed solution first, by the use of implemented information clusterization makes it possible to analyze the image using proposed fuzzy model. In the second stage we use Convolutional Neural Network as a classifier. On the input to the network we use standardized images with all the features forwarded to 3 different kind of layers of the network for pooling, convolution and final classification. As a result we have images with selected objects. Proposed solution has been examined and compared with some similar approaches by the use of open data sets. The results show that developed technique has a good efficiency and high potential for further research and development.

2. Analysis of clusters of points as a basis for detection and data extraction

In this section, we formulate the problem of shape detection in 2D images and describe our method in details. 2D image does not necessarily contain only one object but in many cases presents many of them. Moreover, such an image may contain certain objects that may be only a background, and also include smaller items that should not be detected. To illustrate this situation, let us take a picture, where large items are located in the center of the image, and a variety of patterns is used as a wallpaper on the wall. The purpose of the detector is to find these large components and skip patterns on the wallpaper because it can lead to misclassification. Such assumptions in detection make it a big problem for a complex input.

The proposed method involves the use of filters in order to simplify an input image at the beginning of processing and further calculate the value of a dedicated function, which determines quality of a set of points in a specific area (called cluster). This value is calculated for n equal parts of the image, which serve as measure to choose the most important areas of the input file for final neural classifier. For this operation we have chosen blur and emboss filters, since proposed combination of these two first removes unnecessary details from the image, and after by the emboss we extract the most important object shapes. Proposed fuzzy logic interpreting is introduced to determine if the object in the clus-

Please cite this article as: M. Woźniak, D. Połap, Object detection and recognition via clustered features, Neurocomputing (2018), https://doi.org/10.1016/j.neucom.2018.09.003