ARTICLE IN PRESS
JID: NEUCOM [m5G;March 23, 2018;1:43]
Neurocomputing 000 (2018) 1–15
Contents lists available at ScienceDirect
Neurocomputing
journal homepage: www.elsevier.com/locate/neucom

Correlational examples for convolutional neural networks to detect small impurities

Yue Guo[a,b,*], Yijia He[a,b], Haitao Song[a], Wenhao He[a], Kui Yuan[a]

[a] Institute of Automation, Chinese Academy of Sciences, Beijing, PR China
[b] School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing, PR China

ARTICLE INFO

Article history:
Received 8 May 2017
Revised 22 February 2018
Accepted 6 March 2018
Available online xxx

Communicated by Xiang Bai

Keywords:
Impurity detection
Multi-frame correlation
Convolutional neural network
Correlational example
Sequential training

ABSTRACT

Convolutional neural networks have been significantly improving common object detection performances for a long time. However, targets across frames are independently detected in an image sequence, and object detection methods in multiple frames are generally divided into two main stages: object detection in every single frame and feature map association across frames. In this paper, a multi-frame detection framework is proposed to directly detect small impurities in opaque glass bottles with liquor. Specifically, a convolutional neural network trained with correlational examples simultaneously detects and correlates proposals, and then links them selectively to obtain robust detection results under challenging illuminations. Besides, memory costs of patch pairs become extremely large compared with those of patches, thus a sequential training procedure is introduced to relax hardware requirements. Extensive experiments on impurity datasets demonstrate superior performances of multi-frame detection frameworks with convolutional neural networks than traditional single-frame models.

1. Introduction

Impurity detection and classification in transparent bottles may have already been solved using traditional machine learning methods, and most researchers focus on studying shapes and motions of impurities since backgrounds are relatively visually static. However, to the best of our knowledge, no researches about impurity detection and classification in opaque glass bottles have been published, and performances in our engineering work only based on motions and shapes of impurities have already been proved ineffective because backgrounds in opaque glass bottles are much more dynamic and complex.

Different from impurity detection in transparent bottles, impurities can not be observed outside the bottle wall, and a camera is directly put down into the bottle and samples images above the liquid level, so one of the special characteristics of opaque bottles in our task is that decorative patterns are carved on bottle bottoms. Consequently, considering the locations of cameras and characteristics of bottles themselves, impurity detection performances in an opaque glass bottle depend on several influence factors including non-uniform light exposures of the bottle bottom, bottom fragment motions of the carved pattern, impurity-background discriminative feature extractions and impurity detection methods. Specifically, non-uniform exposures result from various bottom thicknesses and colors of each bottle; fragment motions of the bottom carved pattern are the local shifts of partial designs caused by tilts and fluctuations of liquor level under intense lighting conditions; performances of discriminative feature extractions and detections mainly depend on imaging qualities and impurity detection frameworks.

Impurity detection problem in opaque glass bottles can also be characterized by sampled images: firstly, both impurities have only a few exclusive features; secondly, under the same strength of illumination, ranges of gray images among different bottles look unstable, because sometimes a camera may be put into a nearly transparent bottle, while at other times it may be in a completely dark bottle. Though lighting conditions have been improved with stronger lights, the thicknesses of bottle bottoms vary. Therefore, impurity detection problem in opaque glass bottles still remains challenging.

To address this problem, a multi-frame detection framework with a convolutional neural network is proposed to directly detect and link object proposals in a finite number of frames, and a sequential training method for convolutional neural networks is additionally introduced with restricted memory. Moreover, detection performances in different bottle colors are intuitively analyzed through precision-recall visualization.

The main contributions of this paper are summarized as follows:

* Corresponding author at: Institute of Automation, Chinese Academy of Sciences, Beijing, PR China.
E-mail address: guoyue2013@ia.ac.cn (Y. Guo).

https://doi.org/10.1016/j.neucom.2018.03.017
0925-2312/

Please cite this article as: Y. Guo et al., Correlational examples for convolutional neural networks to detect small impurities, Neurocomputing (2018), https://doi.org/10.1016/j.neucom.2018.03.017

- Only one convolutional neural network trained with correlational examples simultaneously correlates and detects impurities in bottles with liquor under inadequate lighting conditions.
- A multi-frame framework based on a convolutional neural network is applied to link predicted impurities and eliminate background fluctuations.
- Only a small dataset is far from adequate for a convolutional neural network, but after patch balancing and combination, the correlational dataset becomes significantly large. To train and evaluate the convolutional neural network with restricted memory costs, a simple sequential training procedure is introduced to efficiently address this issue.

The rest of this paper is organized as follows. In Section 2, prior works in detection and matching are briefly introduced. In Section 3, a multi-frame detection framework based on a convolutional neural network is proposed, and training networks sequentially with correlational examples is detailed. Analyses of experiments in a single frame and those in multiple frames are independently provided, and then we quantitatively compare multi-frame experiments with single-frame tests in Section 4. Finally, our conclusions and future works are presented in Section 5.

2. Related work

Impurity detection in transparent bottles is the closest application to our works (Section 2.1), and our researches are mainly inspired by many ideas from object detection (Section 2.2), since FPGAs temporarily provide image patches and relevant information, methods based on image patches such as image patch matching (Section 2.3) are further considered but applied differently in our task.

2.1. Impurity detection in transparent bottles

Impurity detection in opaque glass bottles may be the first work in our field since related public works have not been found. Therefore, only the impurity detection in transparent bottles is introduced, which mainly includes following cascade steps:

Initially, moving impurities and backgrounds are different in a continuous image sequence. Second-order difference and accumulative energy were applied to separate moving impurities from backgrounds [1]. Moreover, fused image difference was used to detect motion regions, then fuzzy c-means clustering combined with fuzzy support vector machines divided impurities and backgrounds [2]. Similarly, fuzzy c-means clustering combined with least squares filtering suppressed backgrounds. Next, prior knowledge such as the area, gray values and the location of each proposal region were considered to detect and track impurities [3].

Secondly, moving impurities should be different from bubbles. The maximum length of a moving blob divided by its minimum length was found as the best feature of shape classification, then support vector machines separated bubbles and impurities [1]. Trajectories of moving objects can also be constructed to differentiate bubbles from impurities in ampoule injections [4]. Features including the gray, shape and position of a blob were trained with fuzzy least squares support vector machines to classify impurities and bubbles [3].

Additionally, feature vectors including several shape features are generated to classify different impurities [4].

2.2. Object detection

Many object detection frameworks are applied in a single frame, and the majority of relevant researchers focus on large-scale image detection benchmarks. Mostafa Mehdipour Ghazi et al. [5] used deep neural networks to identify plant species through transfer learning and discovered that simpler models such as AlexNet are easier to be fine-tuned than others like GoogLeNet and VGGNet. However, sometimes datasets are quite different from some specific image detection tasks in special occasions, and finetuning may not even make a network converge. For example, real-time object detection framework, You Only Look Once (YOLO v2) [6], has state-of-the-art performances in common object detection benchmarks, but when applied in our task, it does not even converge if finetuning or training it at the beginning.

Famous benchmarks facilitate the developments and applications of network architectures. Christian Szegedy et al. added 1×1 convolutions into the Inception module to reduce dimensions and achieved the state-of-the-art performance with large-scale datasets [7]. Meanwhile, even with a small number of training data, convolutional neural networks may still perform well. Shiqi Yu et al. applied convolutional neural networks for classification in hyperspectral images, and their work demonstrates that a well-designed network can also have good generalization with few training samples. Specifically, without max pooling layers and fully connected layers, proper convolution kernels and larger dropout rates are used in their network architecture. However, objects occupy large pixel spaces in some public benchmarks such as Imagenet [8], so a trained network naturally prefers to detect big objects, but COCO [9] contains more tiny and occluded objects.

Small object detections might be challenging if directly using a single convolutional neural network, and it may be due to the following reasons: reception fields in high-level layers are quite large, therefore, firstly, they will not encode sufficient informative features if objects are too small; secondly, deeper layers may become less representative for tiny objects when they get more information outside the regions of interest.

On the one hand, to alleviate such problems, object detection tasks can be divided into two steps: locations of relatively notable target contexts and small object detections in these contexts. Junhua Sun et al. [10] proposed an automatic fault recognition system with convolutional neural networks. Specifically, a network in the first stage detects region proposals of side frame keys and shaft bolts, while a network in the second stage identifies faults from those region proposals. Regions of Interest (RoI) pooling operations are applied to extract regions from the last convolutional feature maps.

On the other hand, high-level feature maps can be concatenated with both original images and low-level feature maps to preserve small but informative features. To detect tiny faces, Chenchen Zhu et al. [11] fused high-level feature maps and low-level ones to generate region candidates, and concatenate candidate features after RoI pooling operations in different layers. False positives can be further rejected with the help of human body to detect unconvinced faces. Xiaodan Sui et al. [12] tried convolutional neural networks to segment choroidal boundaries. Specifically, a coarse-scale network is used to learn global features, and a mid-scale network concatenates output feature maps from a model in coarse scale and downsampled original images to get mid-level edge-cost maps, finally, a fine-scale network concatenates original image and upsampled edge-cost maps to refine high-resolution boundaries.

Sampled images may inevitably be polluted by noises in their surroundings, which can lead to meaningless detections, and inspection models can refuse to detect noisy samples beforehand with convolutional neural networks. Honghan Chen et al. [13] proposed a cascaded spatial-temporal deep framework to inspect gastrointestinal tract diseases, a convolutional neural network firstly removes noisy contents seen from a capsule endoscopy, and then another one classifies organs in clear contents.

Features of convolutional neural networks can also assist tracking tasks. Tao He et al. [14] experimented convolutional neural

Please cite this article as: Y. Guo et al., Correlational examples for convolutional neural networks to detect small impurities, Neurocomputing (2018), https://doi.org/10.1016/j.neucom.2018.03.017