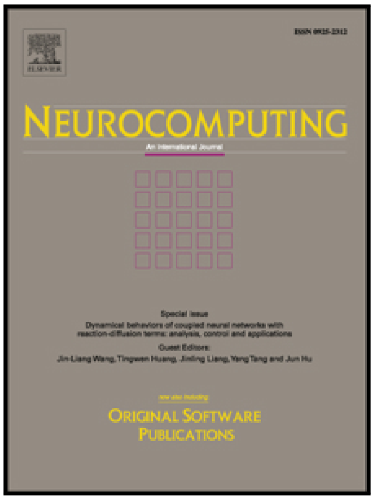
Communicated by Yongzhen Huang

Accepted Manuscript



Deep Learning for Logo Recognition

Simone Bianco, Marco Buzzelli, Davide Mazzini, Raimondo Schettini

PII: S0925-2312(17)30566-0
DOI: 10.1016/j.neucom.2017.03.051
Reference: NEUCOM 18279

To appear in: *Neurocomputing*

Received date: 9 May 2016
Revised date: 2 March 2017
Accepted date: 11 March 2017

Please cite this article as: Simone Bianco, Marco Buzzelli, Davide Mazzini, Raimondo Schettini, Deep Learning for Logo Recognition, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.03.051

Deep Learning for Logo Recognition

Simone Bianco, Marco Buzzelli, Davide Mazzini, Raimondo Schettini

DISCo - Università degli Studi di Milano-Bicocca, 20126 Milano, Italy

Abstract

In this paper we propose a method for logo recognition using deep learning. Our recognition pipeline is composed of a logo region proposal followed by a Convolutional Neural Network (CNN) specifically trained for logo classification, even if they are not precisely localized. Experiments are carried out on the FlickrLogos-32 database, and we evaluate the effect on recognition performance of synthetic versus real data augmentation, and image pre-processing. Moreover, we systematically investigate the benefits of different training choices such as class-balancing, sample-weighting and explicit modeling the background class (i.e. no-logo regions). Experimental results confirm the feasibility of the proposed method, that outperforms the methods in the state of the art.

Keywords: Logo recognition, Deep Learning, Convolutional Neural Network, Data augmentation, FlickrLogos-32

1. Introduction

Logo recognition in images and videos is the key problem in a wide range of applications, such as copyright infringement detection, contextual advertise placement, vehicle logo for intelligent traffic-control systems [1], automated computation of brand-related statistics on social media [2], augmented reality [3], etc.

Traditionally, logo recognition has been addressed with keypoint-based detectors and descriptors [4, 5, 6, 7]. For example Romberg and Lienhart [8] presented a scalable logo recognition technique based on feature bundling, where individual local features are aggregated with features from their spatial neighborhood into Bag of Words (BoW). Romberg et al. [9] exploited a method for encoding and indexing the relative spatial layout of local features detected in the logo images. Based on the analysis of the local features and the composition of basic spatial structures, such as edges and triangles, they derived a quantized representation of the regions in the logos. Revaud et al. [10] introduced a technique to down-weight the score of those noisy logo detections by learning a dedicated burstiness model for the input logo. Boia et al. [11, 12] proposed a smart method to perform both logo localization and recognition using homographic class graphs. They also exploited inverted secondary models to handle inverted colors instances. Recently some works investigating the use of deep learning for logo recognition appeared [13, 14, 15]. Bianco et al. [13] and Eggert et al. [14] investigated the use of pretrained Convolutional Neural Networks (CNN) and synthetically generated data for logo recognition, trying different techniques to deal with the limited amount of training data. Also Iandola et al. [15] investigated a similar approach, proposing and evaluating several network architectures. Oliveira et al. [16] exploited pretrained CNN models and used them as part of a Fast Region-Based Convolutional Networks recognition pipeline. Given the limited amount of training data available for the logo recognition task, all these methods work on networks pretrained on different tasks.

In this paper we propose a method for logo recognition exploiting deep learning. The recognition pipeline is composed by a recall-oriented logo region proposal [17], followed by a Convolutional Neural Network (CNN) specifically trained for logo classification, even if they are not precisely localized. Within this pipeline, we investigate the benefit on the recognition performance of the application of different machine learning techniques in training, such as image pre-processing, class-balancing, sample weighting, and synthetic data augmentation. Furthermore we prove the benefit of adding as positive examples candidate regions coming from the object proposal to the ground truth logos, and the benefit of enlarging the size of the actual dataset with real data augmentation and the use of a background class (i.e. no-logo regions) in training.

2. Proposed Method

The proposed classification pipeline is illustrated in Figure 1. Since logos may appear in any image location with any orientation and scale, and more logos can coexist in the same image, for each image we generate different object proposals, that are regions which are more likely to contain a logo. These proposal are then cropped to a common size to match the input dimensions of the neural network and are propagated through a CNN specifically trained for logo recognition.

In order to have performance as high as possible within this pipeline, we use an object proposal that is highly