Deep Learning Logo Detection with Data Expansion by Synthesising Context

Queen Mary University of London

Hang Su hang.su@qmul.ac.uk
Xiatian Zhu xiatian.zhu@qmul.ac.uk
Shaogang Gong s.gong@qmul.ac.uk

WACV 2017
IEEE 2017 Winter conference on Applications of Computer Vision

1 Introduction

Logo detection challenges:

(1) Small instance size

(2) Non-rigid appearance change

(3) Various designs per logo

1971 1987 1992 2011

(4) Small training data

Dataset	Data Size
ImageNet [1]	1,000,000+
MS COCO [2]	300,000+
BelgaLogos [3]	**1,951**
FlickrLogos-27 [4]	**1,080**
FlickrLogo-32 [5]	**2,240**

Contributions:

(1) A logo data expansion approach by synthesising context

(2) A new logo dataset TopLogo-10 with clothing logos

2 TopLogo-10

Richer context, e.g. person, shoes, clothing, etc.

More logo design variations

Larger logo size variations

3 Data Expansion by Synthesising Context

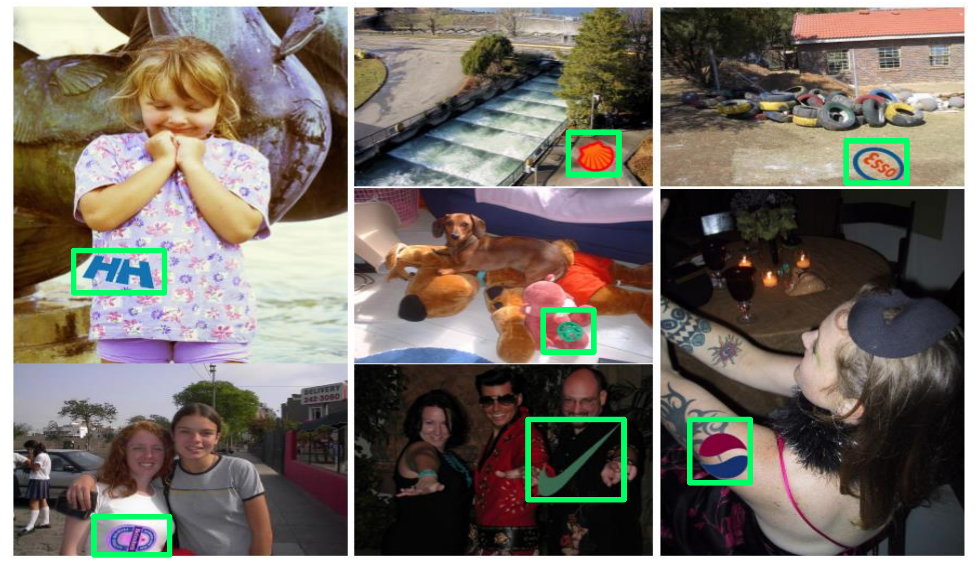
Clean logo exemplar/icons

Synthetic transformation

Scaling Shearing Rotation Colouring

Systhesising context

Synthetic training images with logo in various context

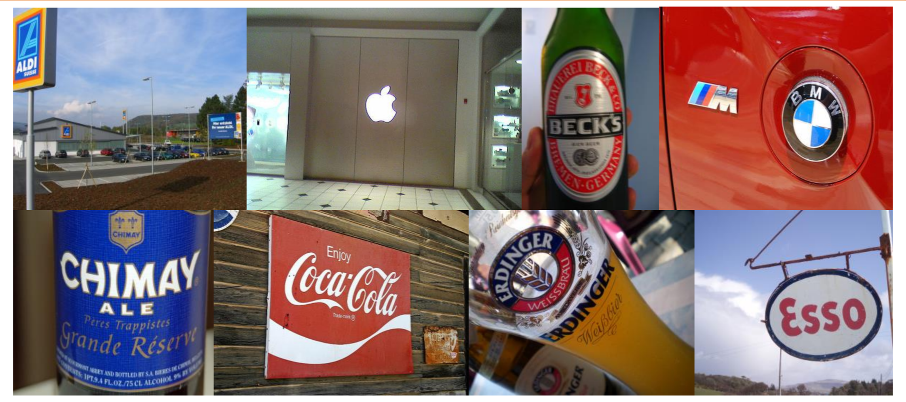


Synthetic Logo Training Dataset Expansion with Enriched Context

Stage 1: Pre-train Model with Synthetic Data

Stage 2: Fine-Tune Model with Real Data

Final Logo Detection Model



Real Logo Image Dataset

Curriculum Model Training:

Easy-to-hard learning strategy

4 Evaluations

Results on FlickrLogo-32

Training Data	mAP
Real data only	50.4
Synthetic data only (32 Cls)	27.6
Synthetic data only (463 Cls)	20.5
Synthetic (32 Cls) + Real	54.8
Synth 463Cls + Real	**55.9**

Significant mAP gain by synthetic pre-training: 5.5% = 55.9 - 50.4
Poor performance by synthetic data only: 27.6% / 20.5% vs 50.4%

Results on TopLogo-10

Training Data	mAP
Real data only	28.5
Synthetic data only (10Cls)	7.3
Synthetic data only (463Cls)	10.2
Synthetic (10Cls) & Real	40.4
Synthetic (463Cls) & Real	**41.8**

Significant mAP gain by synthetic pre-training: 13.3% = 41.8 – 28.5
Poor performance by synthetic data only: 7.3% / 10.2% vs 28.5%

Effect of Curriculum Learning (CL)

Method	FlickrLogo-32	TopLogo-10
Curriculum Learning	**55.9**	**41.8**
Non-Curriculum Learning	30.9	16.3

Significant: 25.0% on FlickrLogo-32 and 25.5% on TopLogo-10 in mAP gain

5 Take-Home Messages

◆ Data expansion by synthetic context is critical to improve deep logo detection given sparse labelled data

◆ Our TopLogo-10 dataset: suggesting that logo detection in diverse real-world scenes is more challenging

References:

[1] J. Deng, etal. ImageNet: A Large-Scale Hierarchical Image Database. CVPR, 2009

[2] Tsung-Yi Lin, et al. Microsoft COCO: Common Objects in Context. ECCV, 2014

[3] A.Joly and O.Buisson. Logo retrieval with a contrario visual query expansion. ICMR 2009

[4] Y. Kalantidis, et a.. Scalable triangulation-based logo recognition. ICMR 2011.

[5] S. Romberg, et al.. Scalable logo recognition in real-world images. ICMR 2011