Computer Vision and Image Understanding 164 (2017) 1–2

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Editorial- Deep Learning for Computer Vision

1. Introduction

Over the past few years computer vision research has made large strides thanks to the advent of deep learning. Starting from breakthrough results in image classification five years ago, there is currently no computer vision problem that has not been affected by this paradigm shift. Our goal has been to capture a snapshot of the fast-paced research that takes place across the broad spectrum of problems at the interface of deep learning and computer vision.

In this context, we are pleased to present a CVIU special issue on Deep Learning for Computer Vision. We have received a total of 34 paper submissions from 19 different countries (Brazil, Canada, China, Czech Republic, Egypt, France, Germany, India, Italy, Japan, Malaysia, Russia, Serbia, Spain, Portugal, Switzerland, Tunisia, UK, USA). The submissions went through an initial check by the guest editors for suitability to the topic, and a few of the submissions were immediately rejected because they were considered o -topic. The remaining papers went through the standard review process with up to three rounds of revisions for some papers. In the end, 12 papers were considered suitable for publication in this special issue.

The accepted publications reflect the exciting research that is currently taking place at the interface of deep learning and computer vision.

In SMC Faster R-CNN: Toward a Scene-Specialized Multi-Object Detector by Ala Mhalla, Thierry Chateau, Houda Maamatou, Sami Gazzah, Najoua Essoukri and Ben Amaraa, the Sequential Monte Carlo (SCM) tech-nique is used to adapt a generic Faster R-CNN detector to a particular camera environment capturing a traffic scene. When combined with spatio-temporal processing this yields substantial improvements over the generic Faster-RCNN baseline.

In Systematic Evaluation of Convolution Neural Network Advances on the ImageNet by Dmytro Mishkin, Jiri Matas and Nikolay Sergievskiy, the authors perform a systematic ablation study of the impact of design choices in deep convolutional neural networks, involving the nonlinearities, pooling functions, learning rates, batch sizes, and normalization options. The results of the paper show that the result of combining the best individual choices yields some-thing only slightly better than the sum of improvements, suggesting that one can explore the impact of every such choices individually.

In Speedup of Deep Learning Ensembles for Semantic Segmentation Using a Model Compression Technique by Andrew Holliday, Mo-hammadamin Barekatain, Johannes Laurmaa, Chetak Kandaswamy, Helmut Prendinger, and Helmut Prendinger, the authors apply model compression to the problem of semantic segmentation, demonstrating that compressed models have real-time performance and similar accuracy to ensembles, compression works on ensembles containing entirely distinct deep architectures, while a good upscaling technique matters more than network depth for semantic segmentation.

In Compact Descriptors for Sketch-based Image Retrieval using a Triplet loss Convolutional Neural Network by Tu Bui, Leonardo Riberio, Moacir Ponti and John Philip Collomosse, the authors present an efficient representation for sketch based image retrieval derived from a convolutional neural network trained with the triplet loss. The learned representation is shown to generalize to novel categories, while the authors also report state-of-the-art retrieval results while indexing based on compact, 56-bit descriptors.

In Deep Compare: A Study on Using Convolutional Neural Net-works to Compare Image Patches by Sergey Zagoruyko and Nikos Ko-modakis, the authors learn a general similarity function for image patches directly from data, exploring Siamese, 2-channel, 2-stream networks and combinations thereof, as well as normalized cross-correlation which are particularly designed for the correspondence task. The authors show that the proposed convolutional networks outperform hand-crafted features by a large margin.

In Hand Pose Estimation through Semi-Supervised and Weakly-Supervised Learning by Natalia Neverova, Christian Wolf, Florian Nebout and Graham Taylor, the authors introduce a hand pose estimation method that relies on direct regression from depth and segmentation maps. Building on this, the authors develop an intermediate representation that allows the adaptation from synthetic to real data, while also developing a restoration-driven method that yields a supervision signal on unlabeled data.

In Weak/semi supervised learning: Harnessing Noisy Web Images for Deep Representation by Phong Dinh Vo, Alexandru Ginsca, Herve Le Borgne and Adrian Popescu, the authors pursue semi-supervised learning approaches to exploit large amounts of unannotated images downloaded from Flickr and Bing. The authors show that one can learn useful representations de-spite the high level of noise in training data, and demonstrate the transferability of the learned representations to new problems.

In Saliency Driven Object Recognition in Egocentric Videos with Deep CNN: toward application in assistance to Neuroprostheses by Philippe Perez de San Roman, Jenny Benois-Pineau, Jean-Philippe Domenger, Florent Paclet, Daniel Cattaert and Aymar de Rugy, the authors propose an

https://doi.org/10.1016/j.cviu.2017.11.006

egocentric framework for real-time object recognition, relying on gaze- fixations recorded by a glass-worn eye-tracker to construct a visual attention signal that guides the object proposal generation. This is shown to substantially improve accuracy, while at the same cutting down computation time to be lower than that required for a visual fixation - thereby facilitating the use of this system in a neuroprosthetic device.

In Hough-CNN: Deep Learning for Segmentation of Deep Brain Regions in MRI and Ultrasound by Fausto Milletari, Seyed-Ahmad Ah-madi, Christine Kroll, Annika Plate, Verena Rozanski, Juliana Maiostre, Johannes Levin, Olaf Dietrich, Birgit Ertl-Wagner, Kai Botze, Nassir Navab and Fausto Milletari, the authors propose a novel volume segmentation approach based on a voting strategy, which is shown to be multi-modal, multi-region, while at the same time encoding priors on anatomical shape and appearance. The method is shown to deliver state-of-the-art results while being registration-free, and scalable to different modalities used in brain region segmentation.

In Improved Gait Recognition Based on Specialized Deep Convolutional Neural Network by Munif Alotaibi and Ausif Mahmood the authors develop a specialized deep convolutional neural network architecture for gait recognition, demonstrating its superiority over subspace-based techniques, while also exploring its potential in an open-set classification setting.

In Multiple Hypothesis Colorization and Its Application to Image Compression by Mohammad Haris Baig and Lorenzo Torresani the authors tackle the multi-modality of the posterior distribution in the colorization task by proposing a tree-structured network to generate multiple plausible colorizations and defining a new learning objective that encourages diversification of the hypotheses. The authors use the output of their CNN to perform image color compression, and show that their proposed method outperforms traditional JPEG color coding by a large margin, producing colors that are nearly indistinguishable from the ground truth at the storage cost of just a few hundred bytes for high-resolution pictures.

In A Study of the Eff ect of Noise and Occlusion on the Accuracy of Convolutional Neural Networks applied to 3D Object Recognition by Jose Garcia Rodriguez, Sergio Orts-Escolano, Sergiu Oprea, Francisco Gomez Donoso, and Miguel Cazorla, the authors study the accuracy of CNNs for 3D object recognition with noise and occlusion, exploring alternative volumetric representations, variants of 2D and 3D CNNs, confirming the robustness delivered by volumetric representations.

We wish to thank CVIU's Editor-in-Chief and Editorial Staff for their encouragement and support that made this Special Issue possible. We want to express in particular our gratitude to our numerous Reviewers who were able to deliver high-quality reviews in a timely fashion.

Ross Girshick
Facebook AI Research
E-mail address: rbg@fb.com

Iasonas Kokkinos
Facebook AI Research and University College London
E-mail address: iasonask@fb.com

Ivan Laptev
Institute National de Recherche en Informatique et Automatique (INRIA)
E-mail address: ivan.laptev@inria.fr

Jitendra Malik
University of Berkeley
E-mail address: malik@eecs.berkeley.edu

George Papandreou
Google
E-mail address: gpapan@google.com

Andrea Vedaldi
University of Oxford
E-mail address: vedaldi@robots.ox.ac.uk

Xiaogang Wang
Chinese University of Hong Kong
E-mail address: xgwang@ee.cuhk.edu.hk

Shuicheng Yan
National University of Singapore
E-mail address: eleyans@nus.edu.sg

Alan Yuille
Jonhs Hopkins University
E-mail address: ayuille1@jhu.edu