

Editorial

Recent Machine Learning Progress in Image Analysis and Understanding

Shengping Zhang ¹, Huiyu Zhou,² and Lei Zhang ³

¹Harbin Institute of Technology, Weihai 264209, China

²Department of Informatics, University of Leicester, LE1 7RH, UK

³University of Pittsburgh, Pittsburgh, USA

Correspondence should be addressed to Shengping Zhang; s.zhang@hit.edu.cn

Received 28 November 2018; Accepted 28 November 2018; Published 10 December 2018

Copyright © 2018 Shengping Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, artificial intelligence and machine learning have attracted increasing attention and achieved great success in both research community and industry especially in the field of multimedia. With the recent progress in machine learning especially in deep learning, many tasks in image analysis and understanding have been applied to solve real problems. For example, since the deep learning based classifier was successfully used in image classification in 2012, deep learning has also been widely used in other computer vision tasks such as video classification and image super-resolution. Learning an effective feature representation from a large number of data is capable of extracting the underlying structure features of the data, which produce better representation than hand-crafted features since the learned features adapt well to the tasks at hand. However, most of the existing deep learning based methods need to learn a huge number of parameters especially with the increasingly complicated network, which restricts their applications in image analysis and understanding in real-time environments.

The primary purpose of this special issue is to organize a collection of recently developed machine learning methods as well as their applications in image analysis and understanding. The special issue is intended to be an international forum for researchers to report the recent developments in this field in an original research paper style. Review articles which describe the current state of the art are also welcomed. From 16 submissions, 7 papers are published in this special issue. Each paper was reviewed by at least two reviewers and revised according to review comments. These papers involve image classification, image segmentation, visual tracking, person reidentification, and so on.

In F. Yang et al.'s paper, the authors implement the MLP neural network on volume data to denoise while preserving the boundary. This method can considerably improve quality of volume data acquired by devices. Then we improve the LH method by combining the regional depth information to achieve the transfer function semiautomatic generation. This method can avoid the influence of noise and make the voxels more centralized. In the LH histogram the voxel distribution at the diagonal line is more concentrated, and the boundary of important objects is effectively emphasized. The features of interest in the data can thus be found exactly by mapping scalar value of boundary voxels which correspond to the points in LH histogram to appropriate opacity and color.

In W. Liu and Y. Zhao's paper, the authors make deep study about the using of word structure of light on the object surface reconstruction. After an image for denoising, they can minimize the impact of other lights to the photographic picture, increasing the compatibility of reading photos; to get each of the height line of the sum and interpolation operations, they then get a smooth three-dimensional reconstruction of the surface of the object. The latter part of the research process will focus on three-dimensional high-precision, high-speed, and real-time reconstruction for further study.

In K. Zhang et al.'s paper, the authors concentrate on identifying tomato leaf disease using deep convolutional neural networks by transfer learning. The utilized networks are based on the pre-trained deep learning models of AlexNet, GoogLeNet, and ResNet. First we compared the relative performance of these networks by using SGD and Adam optimization method, revealing that the ResNet with SGD

optimization method obtains the highest result with the best accuracy 96.51%. Then, the performance evaluation of batch size and number of iterations affecting the transfer learning of the ResNet was conducted. A small batch size 16 combining a moderate number of iterations 4992 is the optimal choice in this work. Our findings suggest that, for a particular task, neither large batch size nor large number of iterations may not improve the accuracy of the target model. The setting of batch size and number of iterations depends on your data set and the utilized network. Next, the best combined model was used to fine-tune the structure. Fine-tuning ResNet layers from 37 to "fc" obtained the highest accuracy 97.28% in identifying tomato leaf disease. Based on the amount of available data, layer-wise fine-tuning may provide a practical way to achieve the best performance of the application at hand. We believe that the results obtained in this work will bring some inspiration to other similar visual recognition problems, and the practical study of this work can be easily extended to other plant leaf disease identification problem.

In M. Sun et al.'s paper, after experimenting with two different data sources using the proposed method, the following conclusions can be drawn: through a large number of experiments, a set of optimal combination parameters suitable for random forest was obtained. That is, we set the number of decision trees at 300 and the number of nodes at 4. When the random forest parameters were optimal, samples of 5% and 10% of the total were selected for the experiment, and the samples of 5% and 10% were added each time. The weighted entropy algorithm was used to select samples with the largest entropy values to train the new training set proposed in this paper. The classifier used the remaining data as the test data to evaluate the performance of the classifier and to test the universality of the classifier. Compared with the traditional classifier based on supervised classification and SVM, we proved via a large number of experiments that the proposed weighted entropy semisupervised ensemble classifier based on random forest showed better classification performance and better universality.

In R. J. Hemalatha et al.'s paper, a method was presented to evaluate the active contour segmentation algorithms to segment the synovial region from arthritis affected finger ultrasound image. Performance analysis metrics like Dice coefficient and Hausdroff distance and statistical analysis metrics like standard error and F-test show the significant difference between the two segmentation method (Caselles and Lankton) for synovial region. Further classification is performed for the derived features such as performance metrics and statistical values. Higher accuracy is described for Lankton as the result of classification process. Hence the output of the research work shows that Lankton method is the best method for synovial region segmentation from ultrasound images.

In W. Liu and H. Wang's paper, the authors propose to use compressed features to model the tracked target's appearance and then use SVM to perform tracking. The experimental results indicate that the proposed method outperforms several state-of-the-art methods. The advantages of the proposed method are twofold: (1) It is good at handling scale changes of the target over time because the used features

are obtained by multiscale wavelet transformation. (2) The speed of the proposed method can achieve real-time because the dimensionality of the used features was reduced by compressed sensing techniques.

M. A. Syed et al.'s paper presents a metric learning approach that exploits both multimodal transforms and cross views impostors to improve the capability of metric to discriminate among different persons as well as enhance rejection capability to decline large number of real world diverse impostors. In real world mostly pedestrian images are multimodal, and in public spaces several persons share similar clothing; therefore, our IRM3 is learned to tackle such issues of reidentification and person tracking in public spaces. Extensive experiments on three challenging datasets (VIPeR, CUHK01, and CUHK03) demonstrate the effectiveness of our IRM3 metric which has outperformed many previous state of the art metrics. In addition, we further intend to extend our approach for testing in real world scenario and intend to solve various other issues for real time implementation.

The accepted papers present recent machine learning progress in image analysis and understanding. We hope that this special issue would attract a major attention of the peers.

Conflicts of Interest

The editors declare that they have no conflicts of interest regarding the publication of this special issue.

Acknowledgments

We would like to express our appreciation to all the authors, reviewers, and the editor-in-chief for great support to make this special issue possible.

*Shengping Zhang
Huiyu Zhou
Lei Zhang*

