

EDITORIAL

Special issue on deep learning based detection and recognition for perceptual tasks with applications

LI-WEI KANG

Deep learning has been popular in artificial intelligence with many applications due to great successes in many perceptual tasks (e.g. object detection, image understanding, and speech recognition). Moreover, deep learning is also critical in data science, especially for big data analytics relying on extracting high-level and complex abstractions as data representations based on a hierarchical learning process. In realizing deep learning, supervised and unsupervised approaches for training deep architectures have been empirically investigated based on the adoption of parallel computing facilities such as GPU or CPU clusters. However, there is still limited understanding of why deep architectures work so well and how to design computationally efficient training algorithms and hardware acceleration techniques.

At the same time, the number of end devices, such as Internet of Things devices, has dramatically increased. These devices usually aim at some deep learning-based perceptual tasks or applications as they are often directly connected to sensors (e.g. cameras) that continuously capture a large quantity of visual data. However, traditional cloud-based infrastructures have been not enough for the demands of the current state of deep learning systems on end devices due to some limitations, such as associated communication costs, latency, security, and privacy concerns induced by current infrastructures. Therefore, the concepts of fog and edge computing have been recently proposed to alleviate these limitations by moving data processing capabilities closer to the network edge.

This special issue focuses on all aspects of deep learning architectures, algorithms, and applications, with particularly emphasizing on exploring recent advances in perceptual applications, hardware acceleration architectures, and deep neural networks over the cloud, fog, edge, and end devices. This special issue has collected five excellent articles reviewed and highly recommended by the editors and reviewers. The first one is “Understanding convolutional neural networks via discriminant feature analysis,” authored by Hao Xu, Yueru Chen, Ruiyuan Lin, and C.-C. Jay Kuo.

This paper addresses the important issue for understanding the convolutional neural networks (CNNs) by analyzing the trained features of a CNN at different convolution layers using two quantitative metrics. The discriminative ability of trained CNN features is finally validated by experimental results. This paper offers important insights into its operational mechanism, including the behavior of trained CNN features and good detection performance of some object classes that were considered difficult in the past.

The second article is “Toward visible and thermal drone monitoring with convolutional neural networks,” authored by Ye Wang, Yueru Chen, Jongmoo Choi, and C.-C. Jay Kuo. This paper presents a visible and thermal drone monitoring system that integrates deep-learning-based detection and tracking modules. Two data augmentation techniques are developed to overcome the problem of the paucity of training drone images especially thermal drone images. As a result, even being trained on synthetic data, the proposed system performs well on real-world drone images with complex background.

The third article is “A deep learning-based method for vehicle licenseplate recognition in natural scene,” authored by Jianzong Wang, Xinhui Liu, Aozhi Liu, and Jing Xiao. This paper proposes a solution to recognize real-world Chinese license plate photographs using the DCNN (deep convolutional neural network)-RNN (recurrent neural network) model. With the implementation of DCNN, the license plate is located and the features of the license plate are extracted after the correction process. Finally, an RNN model is performed to decode the deep features to characters without character segmentation.

The fourth article is “Combining acoustic signals and medical records to improve pathological voice classification,” authored by Shih-Hau Fang, Chi-Te Wang, Ji-Ying Chen, Yu Tsao, and Feng-Chuan Lin. This paper proposes two multimodal frameworks to classify pathological voice samples by combining acoustic signals and medical records. In the first framework, acoustic signals are transformed into static supervectors via Gaussian mixture models. Then, a deep neural network (DNN) combines the supervectors with the medical record and classifies the voice signals. In the second framework, both acoustic features and medical data are processed through first-stage DNNs individually. Then, a second-stage DNN combines the outputs of the first-stage DNNs and performs classification.

National Yunlin University of Science and Technology, Douliu, Yunlin, Taiwan

Corresponding author:

Li-Wei Kang,

Email: lw kang@yuntech.edu.tw

The fifth article is “The artificial intelligence renaissance: deep learning and the road to human-level machine intelligence,” authored by Kar-Han Tan and Boon Pang Lim. This paper looks at recent advances in artificial intelligence, and addresses that a number of problems being considered too challenging just a few years ago can now be solved convincingly by deep neural networks (DNNs). The paper essentially takes a look at the inner working of the representative DNN architectures, highlights a number of recent advancements, and concludes with a discussion of open challenges and opportunities.

Based on the papers published in this special issue, the issue brings the insight of deep learning in different aspects

from theory to application. It is expected that it would be helpful for the readers to better understand how the deep models work and possibly design novel deep networks and applications.

Guest Editors of the special issue:

Dr. Li-Wei Kang (National Taiwan Normal University, Taiwan)

Dr. Wen-Huang Cheng (National Chiao Tung University, Taiwan)

Dr. Yuichi Nakamura (NEC Corp., Japan)

Dr. Jia-Ching Wang (National Central University, Taiwan)