Computer Science and Information Systems 19(3): iii–vi https://doi.org/10.2298/CSIS220300iiiY

Guest Editorial – Real-Time Image processing with deep neural networks and optimization algorithms

Shoulin Yin[1] and Mirjana Ivanović[2]

[1] Software College, Shenyang Normal University
253 Huanghe Bei Dajie, Huanggu District, Liaoning Province, 110034 China
yslin@hit.edu.cn
[2] University of Novi Sad, Faculty of Sciences
Novi Sad, Serbia
mira@dmi.uns.ac.rs

Real-time images are often captured and processed without any buffer delays. Since most real-time images are captured from many sources, the quality of the image resolution may vary. However, due to recent advancements in image processing, there are various types of real-time image processing techniques. Real-time image processing may lead to high computational overhead and delays in the transmission of the images, and to overcome these limitations, deep neural networks techniques (DNNs) and optimization algorithms (OAs) may be an asset moving forward. Deep neural networks (DNNs) approach is very popular due to big data support and automatically features selection, this will reduce the workload of scientists, and also convolution neural networks (CNNs) techniques will be used to increase accuracy as compared to machine learning methods. DNNs, like convolution neural networks (CNN), Deep adversarial network (DAN), long short-term memory (LSTM), autoencoder, and deep belief networks have been used to provide real-time image processing.

Using deep neural networks, various hidden layers within them will capture important features of an image or a frame. When the image is captured on a real-time basis, it can be processed by deep neural networks more efficiently and effectively. However, there may be significant performance pressure on the processing and evaluation of real-time high resolution and multi-resolution images. This special section provides an exemplary forum for researchers to discuss theories and ideas associated with real-time image processing using deep neural networks methods and optimization algorithms. Also, this special section discusses all the challenges and limitations of using deep neural networks models in real-time image processing.

This special section aims to receive high-quality papers that extend the current state of the art with innovative ideas and solutions in the broad area of utilization of deep neural networks in real-time image processing. Contributions may present and solve open research problems, integrate efficient novel solutions, present performance evaluations, and compare new methods with existing solutions. Theoretical as well as experimental studies for typical and newly emerging convergence technologies and use cases enabled by recent advances are encouraged. For this special section potential topics that were suggested for potential authors included but were not limited to the following:

- DNNs/OAs-based real-time image processing techniques
- Intelligent learning algorithms for real-time image reconstruction and processing
- Real-time image security and privacy using DNNs/OAs
- Federated learning methodologies used in real-time image processing

- processing of real-time images in remote sensing applications using DNNs/OAs techniques
- Quality of Experience (QoE) and Quality of Service (QoS) for real-time image processing
- DNNs/OAs pattern recognition in real-time image processing and processing
- Evaluation of enhanced real-time images using DNNs/OAs methods
- Computational-based DNNs/OAs models for detection of abnormalities in real-time captured images
- New objective functions of DNNs/OAs for real-time image reconstruction
- Performance analysis of semantic segmentation of images using DNNs/OAs algorithms
- Limitations of DNNs/OAs and hybrid models for real-time image processing
- Sports and arts image processing by DNNs/OAs algorithms
- Machine learning models, big data/cloud computing/fog computing etc, for real-time image processing

This special section received 42 submissions where the corresponding authors were majorly counted by the deadline for manuscript submission with an open call-for-paper period of 6 months. All these submissions are considered significant in the field, but however, only one-third of them passed the pre-screening by guest editors. The qualified papers then went through double-blinded peer review based on a strict and rigorous review policy. After a totally three-round review, 9 papers were accepted for publication. We believe that this special section brings challenging research papers and novel approaches that will be interesting and useful for readers.

A brief overview to the papers in this section can be revealed below, and we expect the content may draw attention from public readers, and furthermore, prompt the society development.

The first paper entitled “Crowdsourcing Platform for QoE Evaluation for Cloud Multimedia Services” by Asif Ali Laghari et al. presents a novel web-based crowdsourcing platform for the assessment of the subjective and objective quality of experience (QoE) of the video service in the cloud-server environment. The user has the option to enter subjective QoE data for video service by filling out a web questionnaire. The objective QoE data of the cloud-server, network condition, and the user device is automatically captured by the crowdsourcing platform. The proposed system collects both objective and subjective QoE simultaneously in real-time. The paper presents the key technologies used in the development of the platform and describes the functional requirements and design ideas of the system in detail. The system collects real-time comprehensive data to enhance the quality of the user experience to provide a valuable reference.

The second paper entitled “A Novel Motion Recognition Method Based on Improved Two-stream Convolutional Neural Network and Sparse Feature Fusion” by Chen Chen et al. proposes a novel motion recognition method based on an improved two-stream convolutional neural network and sparse feature fusion. In the low-rank space, because sparse features can effectively capture the information of motion objects in the video, meanwhile, they supplement the network input data, in view of the lack of information interaction in the network, they fuse the high-level semantic information and low-level detail information to recognize the motions, which makes the performance of the two-stream convolutional neural network have more advantages.

The third paper entitled "A New Frog Leaping Algorithm-oriented Fully Convolutional Neural Network for Dance Motion Object Saliency Detection" by Yin Lyu et al. proposes a new frog leaping algorithm-oriented fully convolutional neural network for dance motion object saliency detection.

The fourth paper entitled "A Novel Art Gesture Recognition Model Based on Two Channel Region-Based Convolution Neural Network for Explainable Human-computer Interaction Understanding" by Pingping Li et al. proposes a novel gesture recognition based on two channel region-based convolution neural network for explainable human-computer interaction understanding. The input gesture image is extracted through two mutually independent channels. The two channels have convolution kernel with different scales, which can extract the features of different scales in the input image, and then carry out feature fusion at the fully connection layer. Finally, it is classified by the softmax classifier. The two-channel convolutional neural network model is proposed to solve the problem of insufficient feature extraction by the convolution kernel. Experimental results of gesture recognition on public data sets NTU and VIVA show that the proposed algorithm can effectively avoid the over-fitting problem of training models and has higher recognition accuracy and stronger robustness than traditional algorithms.

The fifth paper entitled "Adaptive Wavelet Transform Based on Artificial Fish Swarm Optimization and Fuzzy C-means Method for Noisy Image Segmentation" by Rui Yang et al. proposes a noisy image segmentation method based on FCM wavelet domain feature enhancement. Firstly, the noise image is decomposed by two-dimensional wavelet. Secondly, the edge enhancement of the approximate coefficient is carried out, and the artificial fish swarm (AFS) optimization algorithm is used to process the threshold value of the detail coefficient, and the processed coefficient is reconstructed by wavelet transform. Finally, the reconstructed image is segmented by FCM algorithm.

The sixth paper entitled "BiSeNet-oriented context attention model for image semantic segmentation" by Lin Teng et al. proposes a BiSeNet-oriented context attention model for image semantic segmentation. In the BiSeNet, the spatial path is utilized to extract more low-level features to solve the problem of information loss in deep network layers. Context attention mechanism is used to mine high-level implied semantic features of images. Meanwhile, the focus loss is used as the loss function to improve the final segmentation effect by reducing the internal weighting.

The seventh paper entitled "DRN-SEAM: A Deep Residual Network Based on Squeeze-and-Excitation Attention Mechanism for Motion Recognition in Education" by Xinxiang Hua et al. proposes a residual network based on Squeeze-and-Excitation attention mechanism. Deep residual network is widely used in various fields due to the high recognition accuracy.

The eighth paper entitled "Human Action Recognition Using a Depth Sequence Keyframes Based on Discriminative Collaborative Representation Classifier for Healthcare Analytics" by Yuhang Wang et al. proposes a new deep map sequence feature expression method based on discriminative collaborative representation classifier, which highlights the time sequence of human action features. In this paper, the energy field is established according to the shape and action characteristics of human body to obtain the energy information of human body. Then the energy information is projected onto three orthogonal axes to obtain deep spatial- temporal energy map. Meanwhile, in order to solve the problem of high misclassification probability of similar samples by collaborative representa-

tion classifier (CRC), a discriminative CRC (DCRC) is proposed. The classifier takes into account the influence of all training samples and each kind of samples on the collaborative representation coefficient, it obtains the highly discriminative collaborative representation coefficient, and improves the discriminability of similar samples.

The last paper entitled "A Novel Deep LeNet-5 Convolutional Neural Network Model for Image Recognition" by Jingsi Zhang et al. proposes a novel deep LeNet-5 convolutional neural network model for image recognition. On the basis of Lenet-5 model with the guaranteed recognition rate, the network structure is simplified, and the training speed is improved. Meanwhile, they modify the Logarithmic Rectified Linear Unit (L ReLU) of the activation function. The method showed the better effect.

We would also like to thank Prof. Mirjana Ivanović the editor-in chief of ComSIS, for her support during the preparation of this special section in ComSIS journal. Finally, we gratefully acknowledge all the hard work and enthusiasm of authors, reviewers, and core editorial team without whom the special section would not have been possible.