Guest Editorial: Special Section on Advanced Signal Processing and AI Technologies for Industrial Big Data

POWERED by the Industry 4.0, the modern industry evolves from the traditional manufacturing industry to digital and intelligent industry. Huge amount of complex real-time data are generated from the thousands of industrial sensors in physical and man-made environments. Industrial big data (IBD) afford us an unprecedented opportunity to obtain an in-depth understanding of Internet of Things and facilitate data-driven approaches for industrial optimization and scheduling [item 1) of the Appendix]. However, it is a great challenge for analyzing and processing IBD because of its complex and unstructured characteristic, such as volume, variety, velocity, value, sequence, strong relevance, accuracy and closed loop. The value of IBD can be mined by a series of technologies and methods, including data planning, acquisition, preprocessing, storage, analytical mining, visualization, and intelligent control. The essential goal of analyzing and mining IBD is to discover new patterns and knowledge from complex industrial datasets, and to extract novel valuable information, which promotes product innovation, improves operation level and production operation efficiency of manufacturing enterprises, and expands novel business models.

Graphs offer the ability to model such data and complex interactions among IBD. Recently, graph signal processing defines a framework that allows the extension of classical signal processing concepts and provides a promising way for processing data defined on irregular graph domains [item 2) of the Appendix]. In addition, deep learning has achieved great success on Euclidean data. There are an increasing number of applications where data are generated from the non-Euclidean domain and need to be effectively analyzed. Recently, there is increasing interest in extending deep learning approaches for graph data. The graph neural networks raised by the data scientist can extract the features from the irregular large-scale dataset [item 3) of the Appendix]. Therefore, there has been an emerging trend to take signal processing and artificial intelligence into account when designing effective graph analytic mechanism in industry aggregation areas.

This special section collects the latest ideas and research on advanced signal processing and artificial intelligence (AI) technologies for IBD. Particularly, 14 original articles are accepted and included in the collection on the following pages. The topics of these articles are mainly concerned with deep learning, signal processing, resource allocation for IBD, and so forth. We believe that these articles will play a role in inspiring our readers.

Digital Object Identifier 10.1109/TII.2020.3034627

Summaries of Accepted Articles

The first article "Support Multi-Mode Tensor Machine for Multiple Classification on Industrial Big Data," authored by Ma *et al.*, presents a support multimode tensor machine (SMTM) algorithm by applying the multimode product to generalize the formulation of the STM. Furthermore, the paper presents an efficient algorithm to train the parameters. Experiments conducted on various data sets validate the better performance of SMTM over other algorithms in the multiple classifications and imply the potential of the proposed model for multiple classifications on IBD.

The second article "Reinforcement Learning-Based Multislot Double-threshold Spectrum Sensing With Bayesian Fusion for Industrial Big Spectrum Data," authored by Liu *et al.*, proposes a reinforcement learning-based multislot double-threshold spectrum sensing with Bayesian fusion to sense industrial big spectrum data, which can find required idle channels faster while guaranteeing spectrum sensing performance. They set double thresholds guarantee both high detection probability and spectrum access probability, and propose weighed energy detection to maximize detection probability when the energy statistic falls into the confusion area between the double thresholds.

The third article "Privacy-Assured FogCS: Chaotic Compressive Sensing for Secure Industrial Big Image Data Processing in Fog Computing," authored by Zhang *et al.*, presents a chaotic compressive sensing (CS) scheme for securely processing industrial big image data in the fog computing paradigm, called privacy-assured FogCS. Specially, the Sine Logistic modulation map is used to drive the privacy-assured, authenticated, and block CS for secure image data collection in the sensor nodes. After sampling, the measurements are normalized in the fog nodes. The normalized measurements can achieve the perfect secrecy and their energy values are further masked through the proposed permutation-diffusion architecture.

The fourth article "Data-Driven Deep Learning for Signal Classification in Industrial Cognitive Radio Networks," authored by Liu *et al.*, proposes a novel framework of signal intelligent classification based on deep learning networks in industrial cognitive radio networks. In the proposed framework, they preprocess wireless signals first by Choi–Williams distribution time–frequency analysis and represent by two-dimensional time-frequency images. Then, they extract features of wireless signals through stack hybrid autoencoders.

The fifth article "Low Complexity MIMO-FBMC Sparse Channel Parameter Estimation for Industrial Big Data Communications," authored by Wang *et al.*, proposes a low-complexity sparse adaptive channel estimation (CE) scheme, which is based on a dynamic threshold. This reduces the number of inner product calculations by considering only the columns of the measurement matrix greater than the threshold.

The sixth article "Wide-Attention and Deep-Composite Model for Traffic Flow Prediction in Transportation Cyber-Physical Systems," authored by Zhou *et al.*, proposes a wide-attention and deep-composite model consisting of a wide-attention module and a deep-composite module. In particular, the wide-attention module can extract global key features from traffic flows via a linear model with a self-attention mechanism. The deep-composite module can generalize local key features via a convolutional neural network (CNN) component and a long short-term memory (LSTM) network component.

The seventh article "Task Allocation Strategy for MEC-Enabled IIoTs via Bayesian Network Based Evolutionary Computation," authored by Sun *et al*., proposes a mobile edge computing enabled architecture considering the priority constraints among tasks with the objective to minimize the response time.

The eighth article "A Real-time Defect Detection Method for Digital Signal Processing of Industrial Inspection Applications," authored by Gao *et al.*, proposes a novel defect detection method based on deep learning for digital signal processing of industrial inspection applications. In their method, a module named feature collection and compression network is applied to merge multiscale feature information. Then, a new pooling method, named Gaussian weighted pooling, which provides more precise location information, is used to replace region of interest (ROI) pooling.

The ninth article "Efficient Data Transmission Strategy for IIoTs With Arbitrary Geometrical Array," authored by Zheng *et al.*, develops a system architecture for IBD transmission based on radar-communication integration with an arbitrary geometrical array. They use the manifold separation technique to transform the complex array configuration into regular array and the downlink channel covariance matrix is estimated by exploiting a frequency calibration technique when the uplink channel covariance matrix is received. The computational complexities for the proposed method and other state-of-the-art methods are analyzed.

The tenth article "Variational LSTM Enhanced Anomaly Detection for Industrial Big Data," authored by Zhou *et al.*, proposes a variational LSTM learning model for intelligent anomaly detection based on reconstructed feature representation. An encoder–decoder neural network associated with a variational reparameterization scheme is designed to learn the low-dimensional feature representation from high-dimensional raw data. Three loss functions are defined and quantified to constrain the reconstructed hidden variable into a more explicit and meaningful form. A lightweight estimation network is then fed with the refined feature representation to identify anomalies in IBD.

The eleventh article "A Data-Driven Auto-CNN-LSTM Prediction Model for Lithium-Ion Battery Remaining Useful Life," authored by Ren *et al.*, proposes a new lithium-ion battery remaining useful life (RUL) prediction method, namely auto-CNNLSTM. This method is developed based on a deep CNN and LSTM to mine deeper information in finite data. In this method, an autoencoder is utilized to augment the dimensions of data for more effective training of a CNN and LSTM. In order to obtain a continuous and stable output, a filter is used to smooth the predicted value.

The twelfth article "Intelligent Fault Diagnosis of Rotor-Bearing System Under Varying Working Conditions With Modified Transfer CNN and Thermal Images," authored by Shao *et al.*, proposes a new framework for rotor-bearing system fault diagnosis under varying working conditions by using a modified CNN with transfer learning. First, infrared thermal images are collected and used to characterize the health condition of a rotor-bearing system. Second, a modified CNN is developed by introducing stochastic pooling and leaky rectified linear unit to overcome the training problems in a classical CNN. Finally, parameter transfer is used to enable the source modified CNN to adapt to the target domain, which solves the problem of limited available training data in the target domain.

The thirteenth article "Fatigue Detection With Covariance Manifolds of Electroencephalography in Transportation Industry," authored by Zhang *et al.*, investigates whether the spatial–temporal changes in the relations between electroencephalogram (EEG) channels are specific to different driving states. EEG signals were first partitioned into several segments, and the covariance matrices obtained from each segment were input into a recurrent neural network to extract high-level temporal features. Meanwhile, the covariance matrices of whole signals were leveraged to extract spatial characteristics that were fused with temporal features to obtain comprehensive spatial temporal information.

The fourteenth article "D-twin support regression vector machine (TSVR) Recurrence Prediction Driven by Medical Big Data in Cancer," authored by Yang *et al.*, collects a sample of 50 000 cases of seven cancer patients, including liver cancer, lung cancer, kidney cancer, breast cancer, uterine cancer, stomach cancer, and bowel cancer. They propose a TSVR algorithm based on dependent nearest neighbor weighting, the eplion-TSVR model is improved by a DNN-weighted algorithm with local information mining function, and the solution of the improved model is derived. The prediction accuracy of the model for various cancers can reach more than 91%, which is significantly higher than that of CNN and e-TSVR models.

Liangtian Wan, *Guest Editor*
School of Software
Dalian University of Technology
Dalian 116000, China

Mianxiong Dong, *Guest Editor*
Department of Sciences and Informatics
Muroran Institute of Technology
Muroran 050-8585, Japan

XIANPENG WANG, *Guest Editor*
College of Information and
Communication Engineering
Hainan University
Haikou 570288, China
and also with State Key Laboratory
of Marine Resource Utilization in
South China Sea

GUOAN BI, *Guest Editor*
School of Electrical and Electronic
Engineering
Nanyang Technological University
Singapore 639798

ACKNOWLEDGMENT

The Guest Editors would like to thank all the authors who submitted their valuable contributions and anonymous reviewers who carefully scrutinized and evaluated the articles submitted. They would also like to extend their sincere thanks to the Editor-in-Chief of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, Prof. R. Luo, for providing this opportunity and lots of guidance throughout the process, and the editorial staff for their continuous support and professional instruction during the whole preparation of the special section.

APPENDIX
RELATED WORK

1) Y. Shen and O. Kaynak, "Big data for modern industry: Challenges and trends [point of view]," *Proc. IEEE*, vol. 103, no. 2, pp. 143–146, Feb. 2015.
2) A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, May 2018.
3) X. Zhang and Z. Ge, "Automatic deep extraction of robust dynamic features for industrial big data modeling and soft sensor application," *IEEE Trans. Ind. Informat.*, vol 16, no. 7, pp. 4456–4467, Jul. 2020.