IEEE ACCESS SPECIAL SECTION EDITORIAL: MULTIMEDIA ANALYSIS FOR INTERNET-OF-THINGS

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Digital Object Identifier 10.1109/ACCESS.2019.2915487

EDITORIAL

IEEE ACCESS SPECIAL SECTION EDITORIAL: MULTIMEDIA ANALYSIS FOR INTERNET-OF-THINGS

Big data processing includes both data management and data analytics. The data management step requires efficient cleaning, knowledge extraction, and integration and aggregation methods, whereas Internet-of-Multimedia-Things (IoMT) analysis is based on knowledge modeling and interpretation, which is more often performed by exploiting deep learning architectures. In the past couple of years, merging conventional and deep learning methodologies has exhibited great promise in ingesting multimedia big data, exploring the paradigm of transfer learning, association rule mining, and predictive analytics etc.

Statistics reveal that Internet traffic is shifting from non-multimedia data to multimedia data. This signifies the importance and increase of multimedia usage in our day-to-day activities. Seamless integration, cooperative sensing, connectivity, and autonomy in the IoMT infrastructure open doors to numerous opportunities to improve services and applications through efficient utilization of big multimedia data. However, the heterogeneous nature of big multimedia data demands scalable and customized recommendation frameworks for efficient analysis of big data collected in scenarios like surveillance, retail, telemedicine, traffic monitoring, and disaster management. Recommender systems are the technical response to the fact that we frequently rely on peoples' experiences, cultural norms and regional traditions when confronted with a new field of expertise, where we do not have a broad knowledge of all facts, or where such knowledge would exceed the amount of information humans can cognitively process. This observation in the real world suggests that recommender systems are an intuitive and valuable extension, allowing both end-users and multimedia service providers to take a much more active role in selecting semantically relevant content and providing valuable suggestions. For instance, in smart cities, multimedia sensors allow administrators to actively monitor assets and activities. Improvements in automatic interpretation of multimedia big data can enhance the capacity of smart city administrators by autonomously reacting to emergency situations, and recommending effective actions, thereby reducing response times significantly. Furthermore, novel solutions for multimedia data processing and management in the IoMT ecosystem can enhance quality of life, urban environments, and smart city administration.

Our "Call for Papers" received an enthusiastic response with more than 60 high-quality submissions. Per IEEE ACCESS policy, it was ensured that handling editors did not have any potential conflict of interest with authors of submitted articles. All articles were reviewed by at least two independent referees. The articles were evaluated for their rigor and quality, and also for their relevance to the theme of our Special Section. We considered articles that proposed both solutions tailored particularly for the context of the developing world, and also those that were globally oriented, with solutions that could, by extension, also be applicable in the developing world. After a rigorous review process, we accepted 31 articles to form the Special Section.

In "Exploiting user mobility for time-aware POI recommendation in social networks," the authors, Zheng, *et al.*, propose a novel, hybrid Time-aware POI Recommendation model based on User Mobility, TPR-UM, which improves the POI recommendation accuracy greatly. More specifically, by introducing the implicit region factor, the authors capture the users' mobility through collective user actions and geographical properties of locations. The authors generate the check-in patterns based on different time intervals and different regions to exploit the latent relations between temporal factors and geographical influence.

"Random fuzzy cost-profit equilibrium model for locating a discrete service enterprise," by Jia, *et al.*, proposes a new random fuzzy cost-profit equilibrium model by using uncertainty data and management methods. It presents a hybrid algorithm integrating stochastic fuzzy simulation and particle swarm optimization to solve the location problem of an automobile service enterprise. In addition, since risk factors can impact a decision, this work conducts a risk performance analysis when locating an automotive service enterprise. A practical example is given to illustrate the proposed model and algorithm.

"Multi-scheme frame rate up-conversion using space-time saliency," by Li, *et al.*, proposes a novel frame rate up-conversion algorithm, which adaptively selects multiple motion estimation (ME) schemes to generate absent frames

according to the space-time saliency. The authors first use a space-time saliency classifier to determine which frames are most important to a human observer. Based on the varying saliency over time, the authors adopt high-cost ME scheme on high-salient frame but low-cost ME scheme on low-salient frame.

To improve machine vision in bad weather situations, a multi-class weather classification method is presented based on multiple weather features and supervised learning in ''Multi-traffic scene perception based on supervised learning,'' by Jin, *et al*. First, the underlying visual features are extracted from multi-traffic scene images, and the feature is expressed as an eight-dimension feature matrix. Second, five supervised learning algorithms are used to train classifiers. The analysis shows that the extracted features can accurately describe the image semantics, and the classifiers have high recognition accuracy rate and adaptive ability. The proposed method provides the basis for further enhancing the detection of anterior vehicle during night space time illumination changes, as well as enhancing the driver's field of vision on a foggy day.

In the article ''No-reference stereoimage quality assessment for multimedia analysis towards Internet-of-Things,'' by Jiang, *et al*., a new, no-reference stereoscopic image quality assessment model for multimedia analysis toward Internet of Things is built, based on a deep learning model to learn from class labels and image representations. In this framework, images are represented by natural scene statistics features that are extracted from discrete cosine transform domain, and a regression model is employed to shine upon the quality from the feature vector. The training process of the proposed model contains an unsupervised pre-training phase and a supervised fine-tuning phase, enabling it to generalize over the distortion types and severity. The proposed model shows the correlation with subjective assessment as demonstrated by experiments on the LIVE 3-D Image Quality Database and IVC 3-D Image Quality Database.

''Permeance analysis and calculation of the double-radial rare-earth permanent magnet voltage-stabilizing generation device, by Zhang, *et al*., establishes an equivalent magnetic circuit model of a new double-radial PM generator using the traditional equivalent circuit method and data analysis, introduces the calculation formulas of permeance and leakage permeance (LP) in magnetic circuits in detail, analyzes the influence factors of the LP, and provides the solutions. These studies will have a certain guiding significance for the design and optimization of generator circuits. To obtain the best matching parameters, the main parameters of the generator are tested and analyzed, such as the thickness of PM steel in the magnetization direction, length of PM steel, number of pole pairs, and the thickness of spacer bush. Finally, a prototype is tested, and the results show that the output characteristics of the designed double-radial PM voltage-stabilizing generation device are very good.

A lightweight evaluation system, which can be deployed on common user equipment in commercial mobile networks, is proposed for measuring the user experience of multimedia services in ''A lightweight end-side user experience data collection system for quality evaluation of multimedia communications,'' by Chen, *et al*. The user experience evaluation system can measure the key quality indicators of traditional and emerging services in different scenarios. In contrast to traffic models, this system models typical user behavior to construct complex scenarios of communication networks. In commercial network experiments, the proposed evaluation system achieves stable and efficient performance in complex scenarios, which consist of different types of services and typical user behaviors.

In ''Exploiting convolutional neural networks with deeply local description for remote sensing image classification,'' by Liu, *et al*., the authors explore the uses of local description and feature encoding for deeply convolutional features. Given an input image, the image pyramid is constructed, and different pre-trained CNNs are applied to each image scale to extract convolutional features. Deeply local descriptors can be obtained by concatenating the convolutional features in each spatial position. Hellinger kernel and principal component analysis (PCA) are introduced to improve the distinguishable capabilities of the deeply local descriptors. The Hellinger kernel causes the distance measure to be sensitive to small feature values, and the PCA helps reduce feature redundancy. In addition, two aggregate strategies are proposed to form global image representations from the deeply local descriptors. The first strategy aggregates the descriptors of different CNNs by Fisher encoding, and the second strategy concatenates the Fisher vectors of different CNNs. Experiments on two remote sensing image datasets illustrate that the Hellinger kernel, PCA, and the two aggregate strategies improve classification performance. Moreover, the deeply local descriptors outperform the features extracted from fully connected layers.

''Optimal lightweight material selection for automobile applications considering multi-perspective indices,'' by Pu, *et al*., presents a systematic hierarchical structure of multi-perspective indices for optimal lightweight material selection, including mechanical, durabe, societal, and technical properties. A hybrid evaluation approach (G-TOPSIS) integrating grey relation analysis and techniques for order performance by similarity to ideal solution (TOPSIS) is applied, to evaluate lightweight material alternatives and obtain an optimal one. A case study, i.e., 17 kinds of lightweight materials, is conducted to verify the hierarchical structure and the MCDM method. In addition, a sensitivity analysis is conducted to monitor the robustness of the solution ranking to changes. The results show that this method provides an effective decision-making tool for optimal lightweight material selection for automobile applications.

''Two stage particle filter for nonlinear Bayesian estimation,'' by Wang, *et al*., introduces a two-stage particle filter for nonlinear filtering problem. In the proposed particle filter, each particle will be propagated and updated through

two stages. At time step t, the first stage refers to using the unscented Kalman filtering equations to propagate the particles from time step t−1 in order to obtain the preliminary estimations. Then, at the second stage, the particles will be updated again by the iterated extended Kalman filter to yield the final updated particles. In this way, the estimation accuracy of the particle filter can be improved, which is validated through simulation experiments and real-world application experiments.

In ''CS-CNN: enabling robust and efficient convolutional neural networks inference for Internet-of-Things applications,'' Shen, *et al.*, aim for a step forward in this area. They propose a new compressed CNN model termed CS-CNN for image classification by incorporating the theory of compressive sensing at the input layer of the CNN models to both reduce the resource consumption (evaluated as computation time in this article) and required number of training samples. According to the authors' extensive evaluations on the multiple public data sets for deep learning tasks, e.g., MINST and CIFAR-10, using different metrics, the authors illustrate that the CS-CNN is able to speed up the process of training and inference by a factor of magnitude. Meanwhile, it achieves higher classification accuracy compared with the traditional large CNN models when the size of the training database is small.

In ''Open knowledge accessing method in IoT-based hospital information system for medical record enrichment,'' Xie, *et al.*, propose a novel approach to enrich IoT-based medical records by linking them with the knowledge in linked open data. A case study is conducted on a real-world IoT-based HIS system in association with this approach; the experimental results show that medical records in the local HIS system are significantly enriched and useful for healthcare analysis and decision-making, and further demonstrate the feasibility and effectiveness of this approach for knowledge access.

In ''Convolutional neural networks based fire detection in surveillance videos,'' Muhammad, *et al.*, propose a cost-effective fire detection CNN architecture for surveillance videos. The model is inspired from GoogleNet architecture, considering its reasonable computational complexity and suitability for the intended problem compared to other computationally expensive networks such as AlexNet. To balance the efficiency and accuracy, the model is fine-tuned considering the nature of the target problem and fire data. The experimental results on benchmark fire datasets reveal the effectiveness of the proposed framework and validate its suitability for fire detection in CCTV surveillance systems compared to the state-of-the-art methods.

''General regression neural network and artificial-bee-colony based general regression neural network approaches to the number of end-of-life vehicles in China,'' by Xin, *et al.*, proposes two approaches: a general regression neural network (GRNN) and an optimized GRNN based on an artificial bee colony. These approaches are applied to forecast the number of end-of-life vehicles (ELVs) in China. In addition, the proposed models are used to predict the number of ELVs that will appear in China from 2016 to 2020 by combining the forecasting data for the main factors which influence the number of such vehicles. Theoretical and simulation results indicate that the described approaches are effective and feasible. This article provides practical data support and a better theoretical model for researchers, government managers, and industrial engineers faced with the problems posed by ELVs.

''Energy-efficiency models of sustainable urban transportation structure optimization,'' by Qiang, *et al.,* presents some energy-efficiency models of sustainable urban transportation structure optimization. Their objective functions are to minimize the energy consumption, and their constraints are carbon dioxide emission and traffic efficiency. The models were solved via the artificial fish swarm algorithm and used to optimize the urban transportation structure of Harbin City in China. The results indicate that the models cannot only guarantee the benefit of travelers and reduce the carbon emission, but also minimize energy consumption of urban transportation.

In ''Integrated dynamics control system with ESC and RAS for a distributed electric vehicle,'' by Xie, *et al.*, an integrated control system is developed for improving vehicle handling and stability under critical lateral motions. It includes electric-stability-controller (ESC) and rear-wheel-active-steering (RAS) systems, and coordinates the ESC and RAS controllers based on the β-β phase plane method. The ESC includes a fuzzy logical controller that calculates the yaw moment and an additional rear wheel steering angle based on the vehicle steering states. The RAS control system consists of two parts: a feedback controller and a feedforward controller. When the vehicle is in normal driving situations, the RAS system provides enhanced handling performance. If the vehicle reaches its handling limits, then both ESC and RAS are integrated to ensure vehicle stability. The simulation results demonstrated that the proposed integrated controller system not only can resist external interference and reduce driver fatigue but also can improve both vehicle stability and handling.

In ''Spectral–spatial hyperspectral image classification with K-nearest neighbor and guided filter,'' by Guo, *et al.*, the authors combine k-nearest neighbor (KNN) algorithm with guided filter which can extract spatial context information and denoise the classification results with edge-preserving filtering. To solve the problem of dimension disaster, the authors also take dimensionality reduction into account for HSI classification. To verify the feasibility of the proposed methods, the authors evaluate the performance over four widely used hyperspectral data sets. The experimental results show that with only 5% of samples, the authors' method obtained better performance than improved support vector machine and KNN methods.

In ''Underwater-drone with panoramic camera for automatic fish recognition based on deep learning,'' by Meng, *et al.*, the authors believe that underwater drones

will become a big research topic and find a market in the near future. They developed an underwater drone with a 360° panoramic camera acting as the ''eye'' of the drone. The designs are based on open-source hardware and will be shared as an open source to contribute to the innovation of manufacturing, including drones. The function of the 360° panoramic camera is generated by correcting the images taken by two fisheye lenses. The underwater drone was designed by extending the Raspberry Pi compute module; the frame was designed by OpenSCAD, and the printed circuit board was designed by MakePro. As for the application of the underwater drone, the authors focused on fish recognition for investigating fish species in a natural lake to help protect the original environment.

The article ''Quality management of surveillance multimedia streams via federated SDN controllers in FiWi-IoT integrated deployment environments,'' by Bellavista, *et al.*, originally proposes a model and an architecture that loosely federates FiWi and Edge IoT domains based on the interaction of FiWi and Edge IoT software defined networking controllers; the primary idea is that the federated controllers can seldom exchange monitoring data and control hints the one with the other, thus mutually enhancing their capability of end-to-end quality-aware packet management. To show the applicability and effectiveness of the approach, the original proposal is applied to the notable example of multimedia stream provisioning from surveillance cameras deployed in the Edge IoT domain to both an infrastructure-side server and spontaneously interconnected mobile smartphones; the authors' solution is able to tune the BS behavior of the FiWi domain and to reroute/prioritize traffic in the Edge IoT domain, with the final goal to reduce latency. In addition, the reported application case shows the capability of the solution of joint and coordinated exploitation of resources in FiWi and Edge IoT domains, with performance results that highlight its benefits in terms of efficiency and responsiveness.

In ''Attention alignment multimodal LSTM for fine-gained common space learning,'' by Chen, *et al.*, the authors address the problem with the common space learning approach that maps all related multimodal information into a common space for multimodal data. To establish a fine-grained common space, the aligned relevant local information of different modalities is used to learn a common subspace where the projected fragmented information is further integrated according to intra-modal semantic relationships. Specifically, the authors propose a novel multimodal LSTM with an attention alignment mechanism, namely attention alignment multimodal LSTM (AAM-LSTM), which mainly includes attentional alignment recurrent network (AA-R) and hierarchical multimodal LSTM (HM-LSTM). Different from the traditional methods which operate on the full modal data directly, the proposed model exploits the inter-modal and intra-modal semantic relationships of local information, to jointly establish a uniform representation of multimodal data. Specifically, AA-R automatically captures semantic-aligned local information to learn common subspace without the need of supervised labels, then HM-LSTM leverages the potential relationships of this local information to learn a fine-grained common space.

In ''Fusion of domain-specific and trainable features for gender recognition from face images,'' by Azzopardi, *et al.*, the authors propose a novel approach that fuses domain-specific and trainable features to recognize the gender from face images. In particular, the authors use the SURF descriptors extracted from 51 facial landmarks related to eyes, nose, and mouth as domain-dependent features, and the COSFIRE filters as trainable features. The proposed approach turns out to be very robust with respect to the well-known face variations, including different poses, expressions, and illumination conditions. It achieves state-of-the-art recognition rates on the GENDER-FERET (94.7%) and on the labeled faces in the wild (99.4%) data sets, which are two of the most popular benchmarks for gender recognition. The authors further evaluated the method on a new data set acquired in real scenarios, the UNISA-Public, recently made publicly available.

In ''Towards a generalized approach for deep neural network based event processing for the Internet of Multimedia Things,'' by Aslam, *et al.*, the authors described a generalized approach that can handle Internet of Multimedia Things (IoMT) events as a native event type in event processing engines with high efficiency. The proposed system extends event processing languages with the introduction of operators for multimedia analysis of unstructured events and leverages a deep convolutional neural network based event matcher for processing image events to extract features. Furthermore, the authors show that neural network based object detection models can be further optimized by leveraging subscription constraints to reduce time complexity while maintaining competitive accuracy. The initial results demonstrate the feasibility of a generalized approach toward IoMT-based event processing. The application areas for generalized event processing include traffic management, security, parking, and supervision activities to enhance the quality of life within smart cities.

In ''An improved heuristic optimization algorithm for feature learning based on morphological filtering and its application,'' by Hu, *et al.*, a novel two-stage spectral-spatial HSI classification method is proposed. In the first stage, the standard particle swarm optimization (PSO) is adopted to optimize the parameters, and a novel binary PSO with mutation mechanism is used for feature selection simultaneously. Then, the support vector machine classifier is performed. In the second stage, in order to reduce salt and pepper phenomenon, mathematical morphology post-processing is used to further refine the obtained results of the above stage. Experiments are conducted on two real hyperspectral data sets. The evaluation results show that the proposed approach achieves better accuracy than several existing state-of-the-art methods.

In ''A novel segmentation and representation approach for streaming time series,'' Hu, *et al.*, propose a novel continuous

segmentation and multi-resolution representation approach based on turning points, which subdivides the streaming time series by a set of temporal feature points and represents the time series flexibly. The authors' method cannot only generate more accurate approximation than the state-of-the-art of PLR algorithm, but also represents the streaming time series in a more flexible way to meet the different needs of users. Extensive experiments on different kinds of typical time series datasets have been conducted to demonstrate the superiorities of this method.

''Internet of Things (IoT) for seamless virtual reality space: challenges and perspectives,'' by You, *et al.,* addresses a novel virtual reality (VR) system based on the real world in which we live. The ultimate goal is to implement it as though a VR user freely exists in a place. To this end, it is most important to reconstruct a VR space that provides six degrees-of-freedom (DOF), namely, yaw, pitch, roll, surge, sway, and heave. However, most currently released VR services that are based on the real world limit users' movements to three DOF. Even if the services support six DOF, most are highly complex and based on computer graphics. To overcome this problem, the authors first assume that there is a full Internet of Things (IoT) infrastructure for collecting important data for VR space reconstruction.

''Multi-object tracking by flying cameras based on a forward-backward interaction,'' by Carletti, *et al*., proposes a detection and tracking algorithm based on a novel paradigm suitably combining forward tracking based on local data association with a backward chain, aimed at automatically tuning the operating parameters frame by frame, so as to be totally independent of the visual appearance of the flying area. This also definitively drops any time-consuming manual configuration procedure by a human operator. Although the method is self-configured and requires low-computational resources, its accuracy on a wide data set of real videos demonstrates its applicability in real contexts, even running over embedded platforms. Experimental results are given on a set of 53 videos and more than 60000 frames.

''Query intent recognition based on multi-class features,'' by Qiu, *et al.*, adopts a similarity calculation method based on LSTM and a traditional machine learning method based on multi-feature extraction. It is found that entity features can significantly improve the accuracy of intent classification. Secondly, the accuracy of intention classification based on feature sequences constructed by key entities is up to 94.16% in the field of manual labeling, by using BiLSTM classification model.

In ''Efficient sensitive information classification and topic tracking based on Tibetan web pages,'' Xu, *et al*., present a novel sensitive information classification algorithm and topic tracking algorithm for web page contents. First, a text-sensitive information classification method is proposed based on vector space model and cosine theorem. The main idea is that the different location of sensitive words gives different importance degree at term weight computing. Building sensitive word lists is an artificial work. Compared with sensitive thesaurus, web texts are classified. Sensitive word list is the foundation of classification. After the classification of each text, a new topic tracking algorithm is introduced, which monitors sensitive words during a period of time. The first step is to compute the weight of sensitive words in a fixed period of time and select the top 10 sensitive words. The second step is to select the top 3 sensitive words to track in 10 sensitive words. Experiments show that classification of the text sensitive information is very effective and the result of topic tracking is ideal.

In ''An intelligent system for video surveillance in IoT environments,'' Rego, *et al*., propose an artificial intelligence system for detecting and correcting errors in multimedia transmission in a surveillance IoT environment connected through a SDN. The architecture, algorithm, and messages of the SDN are detailed. The AI system design is described, and the test-bed and the data set are explained. The AI module consists of two different parts. The first one is a classifying part, which detects the type of traffic that is sent through the network. The second part is an estimator that informs the SDN controller on which kind of action should be executed to guarantee the quality of service and quality of experience.

In ''Wireless channel optimization of Internet of Things,'' Yang, *et al*., analyzed the Internet of Things wireless channel, including multichannel analysis, Doppler effect analysis, MIMO channel model analysis and channel correlation characteristics simulation. On this basis, a three-path channel neural network curve approximation model was established, and the CDF and PDF curves of Rayleigh and Rice fading channels were approximated and analyzed. The results showed that the neural network model basically conformed to the reality. According to the 3GPP2 standard, the time domain delay model of the FIR filter is optimized, and the design model has the characteristics of fast calculation speed and high precision. In addition, this article also used the neural network to approximate the amplitude-frequency characteristics of the 64-order fading shaping filter and the 120-order interpolation filter. The simulation results met the requirements.

''Efficient algorithms for mining erasable closed patterns from product datasets,'' by Tuong, *et.al*., first defines erasable closed patterns (ECPs), which can represent the set of EPs without information loss. Then, a theorem for fast determining ECPs based on dPidset structure is proposed and proven. Next, two efficient algorithms [erasable closed pattern mining (ECPat) and dNC_Set based algorithm for erasable closed pattern mining (dNC-ECPM)] for mining ECPs based on this theorem are proposed. The experimental results show that ECPat is the best method for sparse data sets, while dNCECPM algorithm outperforms ECPat algorithm and a modified mining erasable itemsets algorithm in terms of the mining time and memory usage for all remaining data sets.

To conclude, we would like to sincerely thank all the authors for submitting their articles to our Special Section, and the large number of reviewers who kindly volunteered their time and expertise to help us curate a high-quality

Special Section on this important and timely topic. We would also like to thank the former IEEE Access Editor-in-Chief, Professor Michael Pecht, the current Editor-in-Chief, Professor Derek Abbott, and other staff members of IEEE Access for their continuous support and guidance.

ZHIHAN LV, *Guest Editor*
Qingdao University
Qingdao 266071, China

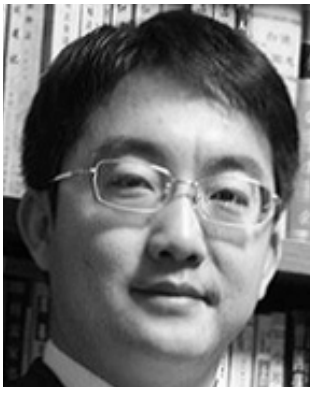
***MARIO VENTO**, Guest Editor*
University of Salerno
84084 Fisciano, Italy

***MINH-SON DAO**, Guest Editor*
Big Data Analytics Laboratory,
National Institute of Information
and Communications Technology (NICT)
Tokyo 184-8795, Japan

***IRFAN MEHMOOD**, Guest Editor*
Sejong University
Seoul 143-747, South Korea

***KAORU OTA**, Guest Editor*
Muroran Institute of Technology
Muroran 050-8585, Japan

***ALESSIA SAGGESE**, Guest Editor*
University of Salerno
84084 Fisciano, Italy



ZHIHAN LV received the Ph.D. degree from Paris 7 University and the Ocean University of China, in 2012.

He has been an Assistant Professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, since 2012. He was a Research Engineer with CNRS, France, a Postdoctoral Research Fellow with Umea University/KTH Royal Institute of Technology, Sweden, and an Experienced Researcher with the Fundación FIVAN, Spain. He was a Marie Curie Fellow with the European Union's Seventh Framework Programme, LANPERCEPT. He is currently an Associate Professor with Qingdao University. He was a Research Associate with the University College London (UCL). He has completed several projects successfully on PC, website, smartphone, and smartglasses. His research interests include multimedia, augmented reality, virtual reality, computer vision, 3D visualization and graphics, serious game, HCI, bigdata, and GIS. His research application fields widely range from everyday life to traditional research fields (i.e., geography, biology, and medicine).

Dr. Lv obtained memberships from several international academic organizations, such as ACM, IEEE, and Eurographics. He is also a Programme Committee Member of the ACM IUI 2015 and 2016, the IEEE BIGDATA4HEALTH Workshop 2016, the IEEE/CIC WIN Workshop 2016, the IIKI2016, and the WASA2016. He has also served as a Reviewer for several journals, such as the IEEE Transactions on Multimedia, the IEEE MultiMedia, *Neurocomputing*, *Computer Networks* (Elsevier), *Telecommunication Systems* (Springer), *Multimedia Tools and Applications*, the *KSII Transactions on Internet and Information Systems*, the *Journal of Medical Internet Research*, *Intelligent Service Robotics*, and *PRESENCE: Teleoperators and Virtual Environments* and conferences, such as the ACM MUM 2013, ACM CHI 2014 and 2015, ACM DIS 2014 and 2016, IEEE InfoVis 2014, ACM UIST 2014 and 2016, ACM MobileHCI 2014–2016, ACM CHIPLAY 2014–2016, ACM SUI 2014 and 2015, ACM ITS 2014 and 2015, IEEE VAST 2014, ACM CSCW 2015 and 2016, IEEE VR 2015 and 2016, ACM IUI 2015 and 2016, IEEE 3DUI 2015 and 2016, ACM TVX 2015, IEEE EuroVis 2015 and 2016, ACM Creativity and Cognition 2015, ACM EICS 2015, ACM IDC 2015 and 2016, IEEE ICSIPA 2015, and GI 2016. He has been an Associate Editor of *PLOS One*, since 2016, the IEEE Access, since 2016, *Neurocomputing*, since 2016, and the *IET Image Processing*, since 2017. He is also the Guest Editor of the IEEE Transactions on Industrial Informatics, *Multimedia Tools and Applications*, *Neurocomputing*, and the *Journal of Intelligent & Fuzzy Systems*.

IRFAN MEHMOOD (M'17) received the B.S. degree in computer science from the FAST National University of Computer and Emerging Sciences, Islamabad, Pakistan, in 2010, and the Ph.D. degree in computer engineering from Sejong University, Seoul, South Korea, in 2015.

He has been involved in IT industry and academia in Pakistan and South Korea for a decade. He is currently an Assistant Professor with the Software Department, Sejong University. His sustained contribution at various research and industry-collaborative projects gives him an extra edge to meet the current challenges faced in the field of multimedia analytics. Specifically, he has made significant contribution in the areas of video summarization, medical image analysis, visual surveillance, information mining, deep learning in industrial applications, and data encryption. He has published over 40 papers in peer-reviewed international journals and conferences, such as *Information Fusion*, *Neurocomputing*, *Sensors*, the *Journal of Visual Communication and Image Representation*, *Multimedia Tools and Applications*, *Computers in Biology and Medicine*, the *Journal of Medical Systems*, *Signal, Image and Video Processing*, *Bio-Medical Materials and Engineering*, the *KSII Transactions on Internet and Information Systems*, NBIS 2015, MITA 2015, and PlatCon 2016. His research interests include the fields of computer vision, image processing, and pattern recognition, with an emphasis on the automated extraction of semantically significant information from images/videos. He is an active interdisciplinary and multidisciplinary researcher, publishing refereed journal and conference papers in the diverse fields of remote sensing, computer vision, algorithm design, and biomedical. He is also an active member of the IEEE. He also provided editorial services in various special issues in top ranked journals, such as *Future Generation Computer Systems* (Elsevier), the *International Journal of Information Management* (Elsevier), the *Journal of Computational Science* (Elsevier), the IEEE ACCESS, *Multimedia Tools and Applications* (Springer), the *IGI Global International Journal of E-Health and Medical Communications* (IJEHMC), and *Computational Intelligence and Neuroscience* (Hindawi). He has been serving as a Professional Reviewer for numerous well-reputed journals, such as the *Journal of Visual Communication and Image Representation*, *Future Generation Computer Systems*, the IEEE ACCESS, *The Journal of Supercomputing*, *Signal, Image and Video Processing*, *Multimedia Tools and Applications*, the *ACM Transactions on Embedded Computing Systems*, and *Enterprise Information Systems*.

MARIO VENTO received the Ph.D. degree in computer engineering from the University of Naples Federico II, in 1989. He is currently a Full Professor of artificial vision, machine learning and cognitive robotics with the University of Salerno, Italy, where he is also the Coordinator of the Artificial Vision Laboratory (MIVIA Lab) and the Dean of the Department of Information and Electrical Engineering and Applied Mathematics. His research interests include real-time video analysis and interpretation for video surveillance applications and affective computing, cognitive robotics, classification techniques, either statistical, syntactic and structural, exact and inexact graph matching, multiexpert classification, and learning methodologies for structural descriptions. He is also a Fellow Scientist of the International Association Pattern Recognition (IAPR). He has served as the Chairman of the IAPR Technical Committee 15 on Graph-Based Representation in Pattern Recognition from 2002 to 2006. He is also an Associate Editor of the *Pattern Recognition* journal.

MINH-SON DAO received the Ph.D. degree in information and communication technology from the International Doctoral School, University of Trento, Italy, in 2005. He was a Scientist with GraphiTech, Italy, from 2005 to 2007. He held a JSPS Post-Doctoral position at the Media Integrated Communication Laboratory (MICL), Graduate School of Engineering, Osaka University, from 2007 to 2010. From 2010 to 2013, he was a member of the MultiMedia Signal Processing and Understanding LAB (mmLAB), University of Trento. From 2013 to 2015, he was a Senior Researcher with the National Institute of Information and Communications Technology (NICT), Japan. He is currently a Principal Lecturer with Universiti Teknologi Brunei (UTB), Brunei, and the Director of the ELEDIA@UTB. His main interests include data mining and big data intelligence analytics for distributed monitoring and event detection, the Internet of Things, and decision support systems in the fields of smart farming, healthcare, and smart cities and communities.

KAORU OTA was born in Aizuwakamatsu, Japan. She received the B.S. and Ph.D. degrees in computer science and engineering from The University of Aizu, Japan, in 2006 and 2012, respectively, and the M.S. degree in computer science from Oklahoma State University, USA, in 2008. From 2010 to 2011, she was a Visiting Scholar with the University of Waterloo, Canada. She was a Japan Society of the Promotion of Science (JSPS) Research Fellow with the Kato-Nishiyama Laboratory, Graduate School of Information Sciences, Tohoku University, Japan, from 2012 to 2013. She is currently an Assistant Professor with the Department of Information and Electronic Engineering, Muroran Institute of Technology, Japan. Her research interests include wireless networks, cloud computing, and cyber–physical systems. He received best paper awards from ICA3PP 2014, GPC 2015, IEEE DASC 2015, IEEE VTC 2016-Fall, and FCST 2017, the 2017 IET Communications Premium Award, and the IEEE ComSoc CSIM Best Conference Paper Award 2018. She was a recipient of the IEEE TCSC Early Career Award 2017. She is also an Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (TVT), the IEEE COMMUNICATIONS LETTERS, *Peer-to-Peer Networking and Applications* (Springer), *Ad Hoc & Sensor Wireless Networks*, the *International Journal of Embedded Systems* (Inderscience), and *Smart Technologies for Emergency Response and Disaster Management* (IGI Global) and the Guest Editor of the *ACM Transactions on Multimedia Computing, Communications, and Applications* (leading), the IEEE INTERNET OF THINGS JOURNAL, the *IEEE Communications Magazine*, the *IEEE Network*, the IEEE WIRELESS COMMUNICATIONS, the IEEE ACCESS, the *IEICE Transactions on Information and Systems*, and *Ad Hoc & Sensor Wireless Networks* (Old City Publishing).

ALESSIA SAGGESE received the Ph.D. degree in computer engineering from the University of Salerno, Italy, and the University of Caen, France, in 2014. She is currently an Assistant Professor with the University of Salerno. Her research interests include basic methodologies and applications in computer vision and pattern recognition. She has been a member of the International Association for Pattern Recognition Technical Committee 15 on Graph-Based Representations in Pattern Recognition, since 2012.