



Deep understanding of big multimedia data

Xiaofeng Zhu¹ · Chong-Yaw Wee² · Minjeong Kim³

Published online: 8 April 2020

© Springer-Verlag London Ltd., part of Springer Nature 2020

The massive increase in the amount of data collected in business, social media, engineering sciences, and computer science makes accessing and analyzing structured, unstructured, and semi-structured data increasingly important [1, 2]. Ability to semantically understand the content of the data can substantially enhance the applications of big data. However, the performance of data understanding is heavily dependent on the learning techniques. In addition to volume, data are naturally comprised of multiple representations or sources in real applications in which multi-source data provide enriched information from different perspectives [3, 4]. Multi-source data understanding is hence one of the most interesting and hottest topics in research and business nowadays and is remarkably useful for practical applications.

Analyzing multi-source big data is a challenging task due to massive volume and multi-source structure of the data [5]. This paper attracts recent studies of multi-source data understanding as well the studies related to this domain to explore efficient and effective solutions of data analytics.

In [6], Yang et al. propose a novel ‘ASTR’ model for end-to-end abstractive review summarization, which leverages the benefits of the supervised deep neural networks, reinforcement learning as well as the unsupervised probabilistic generative model for cross-domain abstractive review summarization. Experiment shows that the

proposed model can generate better sentiment-aware summarization for reviews with different categories and aspects.

In [7], Tan et al. propose a new spectral clustering method based on mutual k-nn. The method uses mutual k-nn to learn the affinity matrix and then employ the affinity matrix to retain local information of the data. Furthermore, the method also utilizes the normalization method to further improve the performance of clustering. Experimental results on eight public data sets verified the effectiveness of the proposed method.

In [8], Du et al. propose that a novel local covariance-based method to solve the problem of the traditional spectral clustering method often ignores the intersection between the different clusters of data. Specifically, the proposed method first learns an initial affinity matrix by adding the local covariance into traditional matrix construction step and then normalizes the obtained affinity matrix to further improve the clustering performance.

In [9], Li et al. propose an unsupervised nonlinear feature selection method via kernel function. Specifically, the method first uses a kernel function to map data into a high-dimensional space, making the data more distinguishable, and then utilize low-rank constraints to remove the influence of noise. Finally, the method employs L_{21} norm conduct feature selection on kernel space. Experiment shows that the method achieves reasonable results on 12 public data sets.

In [10], Zhou et al. propose a low-light enhancement model to solve the problem of imperfect lightness conditions usually lower the visual quality of an image, Specifically, with an input image, the model first generates multiple enhanced images based on a lightness-aware camera response model. These images are then fused at mid-level based on a patch-based image decomposition model. Experimental results show that the proposed model better improves the image quality in terms of visual naturalness and aesthetics.

In [11], Lu et al. propose a novel multitask learning approach called a hybrid representation-learning network

✉ Xiaofeng Zhu
xfzhu0011@hotmail.com

Chong-Yaw Wee
cywee2000@gmail.com

Minjeong Kim
mjkim78@gmail.com

¹ Guangxi Normal University, Guilin, China

² National University of Singapore, Singapore, Singapore

³ University of North Carolina at Greensboro, Greensboro, NC, USA

for text classification tasks. The method consists of two network components: a bidirectional gated recurrent unit with attention network module and a convolutional neural network module. In particular, the attention module allows for the task learning private feature representation in local dependence from training texts and that the convolutional neural network module can learn the global representation on sharing.

In [12], Hu et al. propose an analysis method of students' video watching behavior in MOOCs platform to measure the effectiveness of courses delivered using MOOCs in China Agricultural University. Experimental results show that the proposed method based on Spark platform can quickly and accurately analyze the characteristics of video watching behavior.

In [13], Lu et al. propose a semantic annotation method based on header recognition and data item classification. The method first establishes a description model of the domain to be annotated, and then label the data items on the extracted partial pages and use the header recognition to dynamically expand the domain description model. Finally, the BP neural network algorithm is used to classify the unlabeled data in the header recognition stage, and the unlabeled data items are semantically labeled according to the classification results.

In [14], Zhong et al. propose a novel supervised feature selection method. Specifically, the method uses two subspace learning methods to preserve the local structure and global structure of the data, so as to provide comprehensive information for learning model, and employs adaptive graph learning to adaptively update graph matrix and feature matrix so that the model can output global optimization of spectral feature selection.

In [15], Chen et al. propose a remote sensing image classification algorithm based on the sparse regularized feature learning method. Specifically, after constructing bag of features by using speeded-up robust feature extraction algorithm, direct sparsity optimization-based feature selection method is applied for selecting discriminative features, which is used for constructing support vector machine classifier model.

In [16], Zhu et al. propose a novel region search method to address the RSD problem efficiently. Specifically, the method conducts network expansion search from each region, and upper and lower bounds on spatial-density correlation are defined to prune the search space effectively. In addition, the spatial and density attributes are both considered to prune the search space, so the pruning power is improved in the two domains. Experiments on real and synthetic spatial data sets show the efficiency and effectiveness of the proposed method.

In [17], Liang et al. propose a novel method for detection and recognition of traffic signs. Specifically, in the step

of detection, the method uses the ResNet-50 as the backbone to build pyramidal feature networks, which can greatly enhance the semantic representation of small objects. In the step of recognition, dense blocks are employed in the classification network to boost the accuracy up. Experimental results verified the effectiveness of the proposed method in detecting and identifying traffic signs.

In [18], Liu et al. propose a novel deep convolutional neural network (Res-CNN) to automatically segment acute ischemic stroke lesions from multimodality MRIs. The proposed network employs residual unit to alleviate the degradation problem and use multimodality to exploit the complementary information in MRIs. The experimental results showed that the proposed network outperformed other well established and state-of-the-art segment methods in terms of overlap with the ground truth on both data sets.

In [19], Ma et al. present a model referred as weight-KNN which firstly introduces the CNN feature to address the problem that traditional models only work well with well-designed manual feature representations. Furthermore, the proposed model incorporates a multi-label linear discriminant approach to compute the weighting which improves the accuracy in the subsequent procedures of distance calculation. Experimental results show that the method can achieve a high performance.

In [20], Ouyang et al. propose a new analysis method considering the stability of pulsed inertial neural networks with infinite delay and saturated actuators, which can effectively solve the difficulties caused by unbounded time-varying delay and impulsive effect with saturating actuators. Simulation results show the effectiveness of the new method.

In [21], Chen et al. propose a general spectral sparse regression method to handle abnormal features by learning the joint sparsity and noise features. Specifically, GSSR is conducted in two stages. First, the classic sparse dictionary learning method is used to build the bases of original data. After that, the original data are project to the basis space by learning a new representation via GSSR. Experimental results show that performance of the proposed method.

In [22], Chen et al. propose a multi-task learning-based object detection method by jointly using the color and depth features. The method uses color and depth features to build a pair of background model, forming two streams of the proposed multi-task learning framework, and then generate the final object detection result by fusing the results given by color and depth features.

In [23], Jin et al. propose a sparse Bayesian extreme learning machine (SBELM)-based method for motor imagery EEG classification. The method introduces a sparse Bayesian ELM (SBELM)-based algorithm to improve the

classification performance of MI. Furthermore, SBELM is able to automatically control the model complexity and exclude redundant hidden neurons by combining strengths of both ELM and sparse Bayesian learning.

References

- Shen HT, Liu L, Yang Y, Xing X, Huang Z, Shen F, Hong R (2020) Exploiting subspace relation in semantic labels for cross-modal hashing. *IEEE Trans Knowl Data Eng.* <https://doi.org/10.1109/tkde.2020.2970050>
- Zhu X, Yang J, Zhang C, Zhang S (2019) Efficient utilization of missing data in cost-sensitive learning. *IEEE Trans Knowl Data Eng.* <https://doi.org/10.1109/tkde.2019.29565302019>
- Zhou Y, Tian L, Zhu C, Jin X, Sun Yu (2020) Video coding optimization for virtual reality 360-degree source. *IEEE J Sel Top Signal Process* 14(1):118–129
- Hu R, Zhu X, Zhu Y, Gan J (2019) Robust SVM with adaptive graph learning. *World Wide Web* 1:1. <https://doi.org/10.1007/s11280-019-00766-x>
- Zhu X, Zhu Y, Zheng W (2019) Spectral rotation for deep one-step clustering. *Pattern Recogn* 1:1. <https://doi.org/10.1016/j.patcog.2019.107175>
- Yang M, Qiang Q, Shen Y, Lei K, Zhu J (2018) Cross-domain aspect/sentiment aware abstractive review summarization by combining topic modeling and deep reinforcement learning. *Neural Comput Appl* 1:1–13
- Tan M, Zhang S, Lin W (2018) Mutual kNN based spectral clustering. *Neural Comput Appl* 1:1–8
- Tingting D, Wen G, Cai Z, Zheng W, Tan M, Li Y (2018) Spectral clustering algorithm combining local covariance matrix with normalization. *Neural Comput Appl* 1:1–8
- Li J, Zhang S, Zhang L, Lei C, Zhang J (2018) Unsupervised nonlinear feature selection algorithm via kernel function. *Neural Comput Appl* 1:1–12
- Zhou Z, Feng Z, Liu J, Hao S (2018) Single-image low-light enhancement via generating and fusing multiple sources. *Neural Comput Appl* 1:1–11
- Guangquan L, Gan J, Yin J, Luo Z, Li B, Zhao X (2018) Multi-task learning using a hybrid representation for text classification. *Neural Comput Appl* 1:1–14
- Hui H, Zhang G, Gao W, Wang M (2018) Big data analytics for MOOC video watching behavior based on Spark. *Neural Comput Appl* 1:1–14
- Zhang L, Wang T, Liu Y, Duan Q (2018) A semi-structured information semantic annotation method for Web pages. *Neural Comput Appl* 1:1–11
- Zhong Z (2018) Adaptive graph learning and low-rank constraint for supervised spectral feature selection. *Neural Comput Appl* 1:1–10
- Chen T, Zhao Y, Guo Y (2018) Sparsity-regularized feature selection for multi-class remote sensing image classification. *Neural Comput Appl* 1:1–9
- Zhu S, Wang D, Liu L, Wang Y, Guo D (2018) Inferring region significance by using multi-source spatial data. *Neural Comput Appl* 1:1–9
- Liang Z, Shao J, Zhang D, Gao L (2018) Traffic sign detection and recognition based on pyramidal convolutional networks. *Neural Comput Appl* 1:1–11
- Liu L, Chen S, Zhang F, Fang-Xiang W, Pan Y, Wang J (2018) Deep convolutional neural network for automatically segmenting acute ischemic stroke lesion in multi-modality MRI. *Neural Comput Appl* 1:1–14
- Ma Y, Xie Q, Liu Y, Xiong S (2018) A weighted KNN-based automatic image annotation method. *Neural Comput Appl* 1:1–12
- Ouyang D, Shao J, Cheng H (2018) Stability property of impulsive inertial neural networks with unbounded time delay and saturating actuators. *Neur Comput Appl* 1:1–10
- Chen T, Guo Y, Hao S (2018) Unsupervised feature selection based on joint spectral learning and general sparse regression. *Neur Comput Appl* 1:1–9
- Chen Z, Wang X, Fan T, Lizhong X (2018) Color–depth multi-task learning for object detection in haze. *Neural Comput Appl* 1:1–9
- Jin Z, Zhou G, Gao D, Zhang Yu (2018) EEG classification using sparse Bayesian extreme learning machine for brain–computer interface. *Neural Comput Appl* 1:1–9

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.