IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received November 25, 2018, accepted December 6, 2018, date of publication December 12, 2018, date of current version January 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2886366

A Novel Hot Topic Detection Framework With Integration of Image and Short Text Information From Twitter

CHENGDE ZHANG[1], SHAOZHEN LU[1], CHENGMING ZHANG[1], XIA XIAO[2], QIAN WANG[1], AND GAO CHEN[3]

[1]School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan 430073, China
[2]Laboratory Management Center, Wuhan Qingchuan University, Wuhan 430204, China
[3]School of Electrical Engineering and Intelligentization, Dongguan University of Technology, Dongguan 523808, China

Corresponding author: Chengde Zhang (chengdezhang@zuel.edu.cn)

ABSTRACT Twitter exhibits several characteristics, including a limited number of features and noisy text information. Extracting valuable information from Twitter has made hot topic detection a challenging task. In this paper, a novel four-stage framework is proposed to improve the performance of topic detection. Data preprocessing is the first stage. Deep learning is then exploited to enrich short text information via image understanding. Next, improved latent Dirichlet allocation is used to optimize the image effective word pairs, which improves the accuracy of the extracted topic words. Finally, both short text and images are integrated for topic detection, in which the corresponding topics are mined based on fuzzy matching of topic words. A large number of experiments show that the proposed framework significantly improves the performance of topic detection and outperforms the selected baseline methods.

INDEX TERMS Deep learning, LDA, topic detection, Twitter.

I. INTRODUCTION

A topic is a seminal event or activity, along with all directly related events and activities [1]. Therefore, we can infer that a topic consists of events and activities. A ‘‘hot topic’’ is defined as a topic that appears frequently over a period of time. Generally, a ‘‘hot topic’’ has the following characteristics: (1) it appears in many news stories on one or more news channels; (2) it has strong continuity, which means that many different events relevant to the topic are also reported; and (3) its popularity changes over time.

Due to the obvious features of publishing and receiving messages and the rapid spread of communication, many people regard Twitter as the main channel for event exposure and information sharing. According to Twitter's second-quarter financial statements released on June 30, 2018, the average number of monthly active users on Twitter is 335 million, and the average monthly active users on Twitter in the United States is 68 million. Twitter has a large influence on social media with its billions of monthly active people. How to quickly and accurately discover hot topics from Twitter is a challenging task.

Concise and limited Twitter information significantly affects the topic detection effect. Considerable work has been performed to improve the topic detection effect by improving the topic model [2], [3] from the original vector space model (VSM) to the LDA model. However, there has been little improvement in the retrieval effect. Additionally, short text information has been extended via combination with the external semantic knowledge base [4]. Therefore, it has become crucial to be able to automatically group relevant tweets together. Both images and text are useful information for topic detection. Generally, images on Twitter have a great impact on retweets and are more likely to attract people's attention. Thus, the image content occupies a large proportion of the content on Twitter. Different types of data are semantically related to the same tweet, so complementary enhancements can be achieved when they present the same information. It is believed that the integration of both images and short text information can enrich short text information to improve the performance of topic detection.

In this paper, a novel framework that integrates image and short text information for topic detection is proposed. First, deep learning is explored to enrich short text information via image understanding. Second, improved LDA is used to optimize the image effective word pairs. Finally, the information

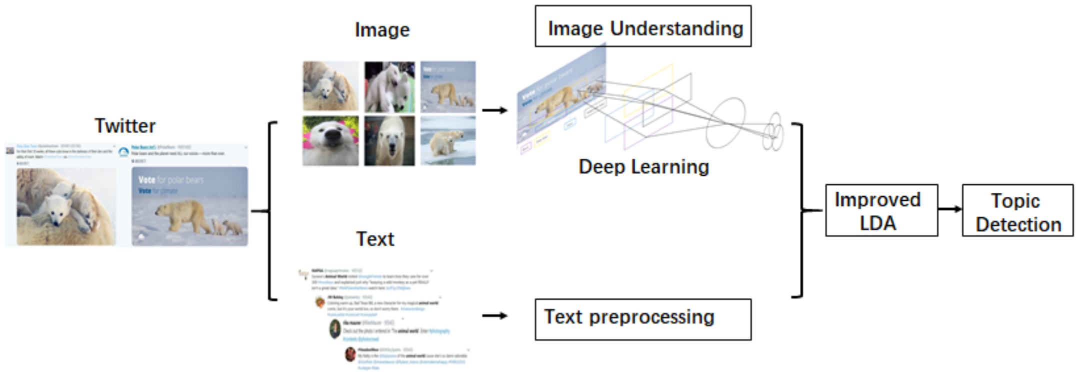


FIGURE 1. Hot topic detection framework with integration of image and short text information from twitter.

from these two steps is integrated for topic detection via fuzzy matching of topic words.

The main novelty and contributions of this paper are summarized as follows:

1) A novel framework that integrates text and image information to improve the performance of topic detection is proposed.
2) We apply deep learning to image level topic detection to enrich short text information by bridging the gap between images and high-level semantic concepts.
3) LDA is improved by optimizing the image effective word pair sets, which improves the accuracy of extracted topic words.

The rest of this paper is organized as follows. Section 2 presents the relevant work. The framework for topic detection is discussed in Section 3. The experiments and results are presented in Section 4. Section 5 concludes this paper.

II. RELATED WORK

A. TOPIC DETECTION

The goal of topic detection is to detect new topics and track the known events in text news streams. Many studies in TDT have been performed [5]–[12]. The VSM converts the text into a space vector and uses the term frequency-inverse document frequency (TF-IDF) function to calculate the feature weight of the word to obtain topics [4]. To detect the multiword relationship in the document and improve the missing semantics of the VSM, latent semantic analysis (LSA) was proposed [13]. To address the gap in the polysemy of LSA, probability latent semantic analysis (PLSA) was proposed [14]. PLSA uses the expectation-maximization (EM) algorithm to obtain the generation probability of each word in the document. At the same time, the Bayesian framework was integrated on the basis of PLSA, and the LDA model was obtained [15]. LDA uses the topic-word matrix to represent the document as the probability distribution of the topic and then mines the specified number of topics from the document set, overcoming the shortcomings of the PLSA model parameters and easy overfitting. An expanded training set is used to enrich the semantic features of the text information. For example, Blei *et al.* [16] combined the LDA model with the time series features of the report for incremental topic detection, reducing the complexity of topic detection. Guo [17] introduced text tag information into LDA, which improved the readability of the theme. Other scholars improved the accuracy of topic mining by improving the LDA model. For example, the hierarchical Dirichlet process (HDP) was proposed to overcome the shortcomings of manually determining the number of topics [18]. It can accurately estimate the distribution parameters of the document set.

B. IMAGE UNDERSTANDING

Image understanding is a technique in which a computer processes, analyzes, and understands images to identify targets and objects in various modes. Recently, many people have begun to study image recognition based on semantic information understanding. In terms of semantic representation, AHP is used to extract image semantic features [19], [20]. Then, a model based on image semantic understanding is realized, which makes the mapping of low-level features and high-level structures more convenient, and its use can effectively avoid the problem of image segmentation in semantic representation. In the process of semantic extraction, a weak segmentation semantic extraction method was proposed [21], and a 2D hidden Markov model was used to extract the semantic features of images [22]. In 2006, deep learning was proposed to win the 2012 ImageNet image classification competition [23]–[25], which started an increased interest in deep learning in the field of computer vision. Deep residual learning was proposed to reduce the error rate on the ImageNet test set to 3.57%.

Motivated by the functionality of deep learning, we explore the utilization of deep learning-based image understanding to obtain more effective information. Compared with noisy terms, which may appear across a long range, images in topics generally appear in a relatively small range and are less noisy. This motivates the study of using image information to improve the performance of topic detection in this paper.

III. PROPOSED FRAMEWORK

The proposed framework is illustrated in Figure 1. As seen from the figure, this framework consists of four stages (enclosed in the rectangular boxes), namely, text

TABLE 1. Process of generating document d.

Input: V matrix, β; the Dirichlet topic prior, α; **Output**: document, $\{w_1, w_2, \cdots, w_N\}$;
Generate N $\backsim$ Poisson(ξ); Generate $\theta_d \backsim Dir(\alpha)$; **For** each i in $w_1, w_2, \cdots, w_N$ **do** Generate topic $z_i Multinomial(\theta_d)$; Generate word w_i from $p(w_i\|z_i, \beta)$, a multinomial probability conditioned on the topic z_i; **Return** $w_1, w_2, \cdots, w_N$;

preprocessing, image understanding, improved LDA, and topic detection.

In the text preprocessing stage, terms extracted from tweets are treated as textual features. Due to noisy user-supplied tag information, text words are pruned via word stemming, special character removal, and so on.

In the image understanding stage, deep learning is applied to interpret the image content as analyzable text semantic information combined with short text, and thus, the semantic space is accurately enriched.

In the improved LDA stage, LDA is improved by optimizing the image effective word pairs, which improves the accuracy of the extracted topic words.

In the topic detection stage, we detect the corresponding topic content based on the fuzzy matching of the topic words.

A. IMAGE UNDERSTANDING BASED ON DEEP LEARNING

Figure 2 is a structural diagram of the neutral image caption (NIC); the generation of the image description is mainly composed of three steps.

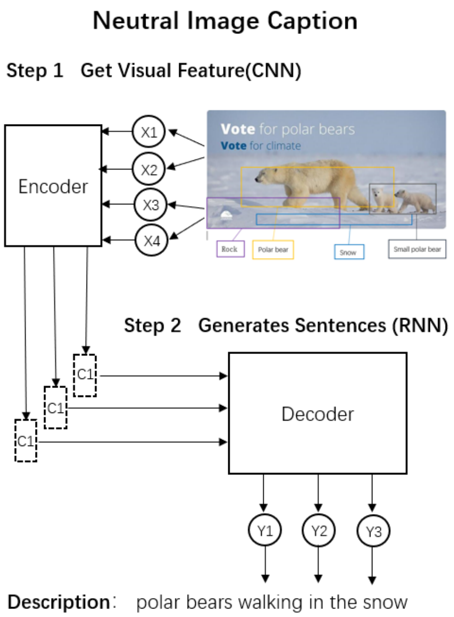


FIGURE 2. Neutral image caption structure diagram.

First, we identify the graphic features using a vision Convolutional Neural Networks(CNN). Images and words are mapped to the same space; images are mapped through the CNN, and the words are mapped using word embedding. We finally obtain the specific meaning of each object in the image. The CNN can embed the input image in a fixed-length vector, so we pretrain a CNN image classification task to obtain the image encoder.

In the second step of our model, decoding is achieved through a long-short term memory (LSTM) network. In this model, the LSTM network is trained as a language model based on image coding. The model utilizes a looping neural network that encodes a variable length input into a fixed dimensional vector and uses this representation to "decode" the vector to the ideal output statement.

$$\theta^* = argmax \sum_{(I,S,\theta)}^{n} \log p(S|I;\theta) \tag{1}$$

Third, we use this formula to calculate the maximum probability of correctly describing the image, so we can obtain the final image description information under this standard. θ is a parameter of the model, I refers to an image, and S is the corresponding translation result.

In Experiment 4.1, We crawled thousands of tweets from ten topics on Twitter including both text and image information. After dataset training, the deep learning model was used to recognition image features. Semantic text information of the image was obtained by deep learning and then merged with text information to enrich the short text semantic space.

B. IMPROVED LDA

LDA is a generation probability model. It can be broken down into a document layer, a topic layer, and a vocabulary layer. The integration of image descriptions alleviates the data sparseness of the document layer and enriches the semantic content of the lexical layer. Meanwhile, LDA assumes that each document consists of a mixture of multiple topics, and each word is generated by a randomly determined topic. Given a corpus D containing M documents and V different words, the process of generating document d is shown in Table 1.

Both α and β are corpus-level parameters, and θ is a variable of the document layer; z_i and w_i are variables of the lexical layer, and N is the number of words in document d. For given α and β, the joint distribution of document d's mixed topic ratio θ_d, the set of topics $z = \{z_1, \cdots, z_N\}$, and the set of words $w = \{w_1, \cdots, w_N\}$ can be expressed by formula (2):

$$p(\theta_d, z, w|\alpha, \beta) = p(\theta_d|\alpha)\prod_{i=1}^{N} p(z_i|\theta_d)p(w_i|z_i, \beta) \quad (2)$$

where α is the k-dimensional vector, β is a matrix of V, K is the number of topics, and V is the number of different words in the corpus. w_i is the $i-th$ word in document d, and z_i is the topic of the $i-th$ word of document d. $p(z_i|\theta_d)$ calculates the probability that the $i-th$ word's topic is z_i of document d, and $p(w_i|z_i, \beta)$ is used to represent the probability of choosing w_i based on z_i and topic-word matrix β; by multiplying the two and multiplying by the probability of θ_d determined by the prior distribution α, the joint distribution of θ_d, z and w can be calculated. Based on formula (2), the probability of generating document d can be expressed by formula (3):

$$p(d|\alpha, \beta) = \int p(\theta_d|\alpha)(\prod_{i=1}^{N} p(z_i|\theta_d)p(w_i|z_i, \beta))d \quad (3)$$

where α and β are learned from the corpus by LDA. After the LDA training process, a topic-word matrix is generated, which records $P(w|z_{1:K})$. For each new document, the topic-word matrix can be used to infer the most likely topic for this document.

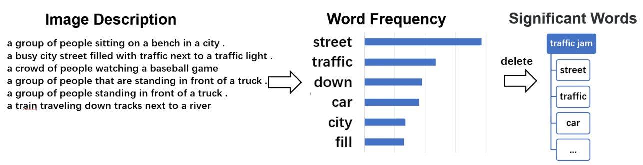


FIGURE 3. Image description processing.

As shown in Figure 3, to reduce the noise of the image content, we propose a new method. First, a large number of images related to the keyword are searched. Then, the images are described, and a descriptive text that records each image is generated. Additionally, after calculating the frequency of the word pairs' occurrence, we define a standard x. If the frequency is higher than x, the word pair is chosen. Next, common words and stay words are deleted, and the valid image description word pairs are retained. Finally, they are merged with the text information. The valid word pairs of keywords for the image descriptions are shown in Figure 4.

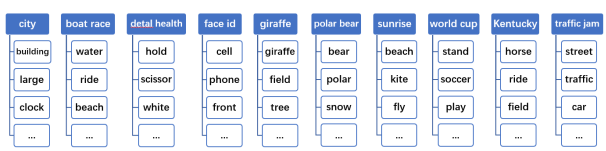


FIGURE 4. Valid word pairs for topics.

By filtering valid word pairs in the image description, not only is the rich semantic information retained, but also, many noise words are eliminated.

C. TOPIC DETECTION

The Levenshtein distance (Lev) is used to measure similarity. The similarity value (referred to as sim) is calculated by formula (4).

$$sim = \frac{Max(l_x, l_y) - Lev}{Max(l_x, l_y)} \quad (4)$$

where l_x and l_y are the lengths of string x and y; $Max(x, y)$ is the maximum of l_x and l_y; and Lev is the value of the $matrix[l_x][l_y]$, which can be expressed by formula (5).

$$matrix(l_x, l_y) = \begin{cases} max(l_x, l_y) if min(l_x, l_y) = 0, \\ min \begin{cases} matrix(l_x - 1, l_y) + 1 \\ matrix(l_x, l_y - 1) + 1 \\ (matrix(l_x - 1, l_y - 1) + 1)_{(l_x \neq l_y)} \end{cases} \end{cases} \quad (5)$$

TABLE 2. Topic clustering.

Input: topic-word bags, $B_1, \cdots, B$; similarity threshold, S; corpus, M;
Output: clusters of $B_i, i \in [1,]$, $C_1, \cdots, C_K$;
Initial clusters of $B_i, i \in [1,]$;
For $i = 1; i < K; i++$ **do**
For each of the M corpus d **do**
$sim_{B_i,d} = Fuzzywuzzy(B_i, d)$;
if $sim_{B_i,d} > S$ **then**
classify d **into** C_i;
Return $\{C_1, \cdots, C_K\}$;

As shown in Table 2, the input is the K topic word bags obtained by LDA, the similarity threshold S, and the corpus M, and the output is K clusters. First, initialize the cluster with $\{B_1, \cdots, B_K\}$ as the centroid. Second, for each B_i, $i \in [1, K]$, calculate its sim to each document d in M using formula (4), and if $sim > S$, classify B_i and d into one class. Finally, K clusters $\{C_1, \cdots, C_K\}$ with $\{B_1, \cdots, B_K\}$ as the centroid are obtained. The *F-measure* is used to evaluate the effect of the topic detection.

IV. EXPERIMENT

A. DATASET

To evaluate the performance of the proposed method, we crawled 10 hot topics, which consist of 34,554 tweets with 733,861 words, 22,647 images and 244,149 words in the image descriptions based on deep learning. These tweets and images were collected from Twitter and selected based on the 10 hot topics in twitter's historic top. These tweets covered different characteristics of topics with rich image information. More details are shown in Table 3.

B. PERFORMANCE EVALUATION

The *precision (P)*, *recall (R)*, and *F1 measure (F1)* are used to evaluate the performance of topic detection, as defined in formulas (6), (7) and (8).

$$Precision = \frac{B_i^+}{A_i} \quad (6)$$

TABLE 3. Dataset information.

ID	Topic	Tweets	Images	Words	Words of image description
1	City	2,425	1,063	55,781	20,263
2	Boat Race	1,292	1,757	55,781	7,561
3	Dental Health	2,984	1,348	61,597	15,219
4	Face Id	4,612	1,708	147,256	18,702
5	Giraffe	3,950	3,085	68,668	33,735
6	Polar Bear	2,202	1,982	53,963	18,864
7	Sunrise	5,319	1,681	75,598	55,769
8	World Cup	3,949	5,369	69,705	20,581
9	Kentucky	2,894	2,810	51,976	21,614
10	Traffic Jam	4,927	1,844	93,536	31,841
Total		34,554	22,647	733,861	244,149

$$Recall = \frac{B_i^+}{B_i} \tag{7}$$

$$F1 = \frac{2PR}{(P+R)} \tag{8}$$

where B_i^+ is the number of correctly grouped positive tweets for cluster A_i, and B_i is the number of positive samples in the ground truth. "LDA-Text" and "LDA-Text+Image" are used as the baselines. The performance comparison is presented in Figures 5, 6, and 7. Due to the tradeoff between *precision* and *recall*, *F1*, which considers both *precision* and *recall*, is introduced, and we focus on the *F1 measure*. It can be easily seen that our proposed framework outperforms the baseline methods, with significant improvements.

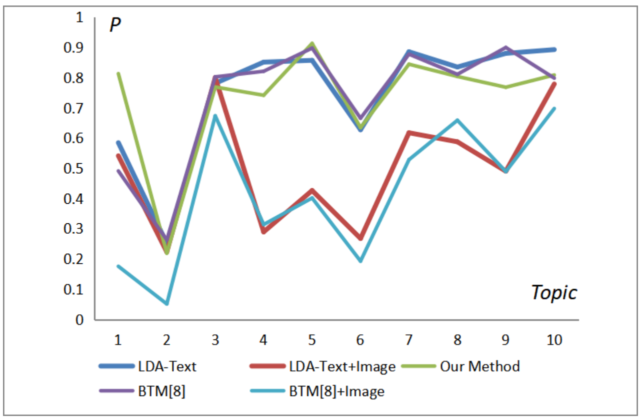


FIGURE 5. Performance comparison - P value.

As shown in Figure 5, overall, we can see that the performances of the methods using image information are poor, as the image description information adds many invalid words. It can be found that the P value of "LDA-Text+Image" and "BTM[8]+Image" are almost always lower than that of "LDA-Text". Taking topic 4 as an example, the P value of "LDA-Text" is 0.85, while the P value of "LDA-Text+Image" and "BTM[8]+Image" are 0.29 and 0.32, respectively. Our method achieves a similar P value as the "LDA-Text" method, as the improved LDA improves the relevance within topics through relevant word pairs, with a significant reduction in the proportion of the correct topics.

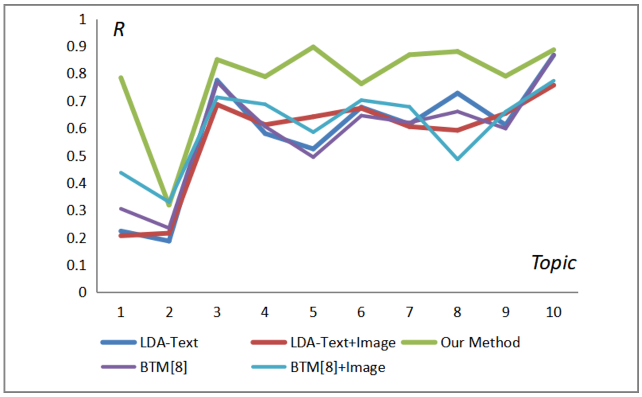


FIGURE 6. Performance comparison - R-value.

As shown in Figure 6, overall, it can be seen that our proposed method achieves promising results compared to all the other baseline methods. We further observe that the recall values for the proposed framework are almost always higher than those of all the other classifiers. This encouraging observation demonstrates that the proposed framework has the ability to help topic detection without missing too many tweets. The method of "LDA-Text+Image" did not achieve as significant an improvement in the R-value due to the introduction of many invalid word pairs. However, for the improved LDA model, the R-value of the processed image description and text fusion is higher than that of the other modes. For example, the recall value of topic 5 increases from 0.51 to 0.9. The reason is that on the basis of the original short text, we add the valid word pairs of the image description, and the frequency of these words appearing is very high in this topic. Thus, the probability of the topic being retrieved greatly increases, and the R-value obviously improves.

As shown in Figure 7, it can be easily seen that the *F1* values significantly improve. More encouraging, the best results reach 90%. It can be inferred from the results that the image content is full of useful information for topic detection. Therefore, the images are good cues for mining tweets that are more relevant to the topics. In contrast, short texts are relatively general, broad, and noisy. Even though the image description may create more noisy information, the image and short text information can complement each other.

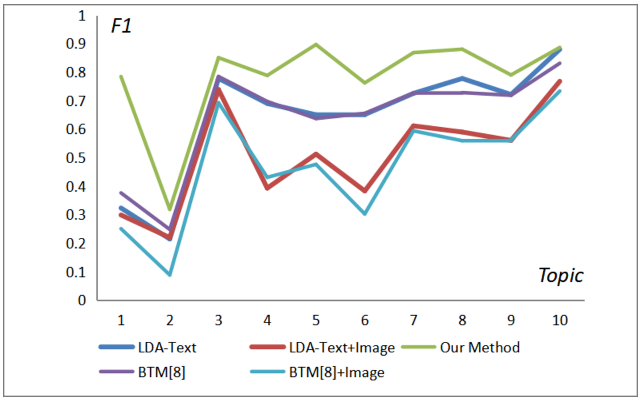


FIGURE 7. **Performance comparison - F value.**

Ultimately, topic detection with short text and images can identify more related tweets. That is, our proposed framework can group more positive information without misclassifying too many negative tweets.

V. CONCLUSION

Considering the very large number of tweets, it is a painstaking task to explore hot topics. Due to the unique characteristics of tweets, such as the limited number of words and noisy text information, topic detection has become a challenging task. In this paper, we proposed a novel 4-stage topic detection framework that integrates short text and image information, aiming to improve the performance of topic detection. Deep learning is applied to image understanding to interpret image content as analyzable text semantic information to enrich short text information. In addition to the limited and noisy short text and images of tweets, we will try to incorporate news websites to obtain more useful information in future work.

REFERENCES

[1] K.-Y. Chen, L. Luesukprasert, and S. C. T. Chou, "Hot topic extraction based on timeline analysis and multidimensional sentence modeling," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 8, pp. 1016–1025, Aug. 2007.

[2] H. Shi, J. Mao, T. Xiao, Y. Jiang, and J. Sun, "Learning visually-grounded semantics from contrastive adversarial samples," in *Proc. 27th Int. Conf. Comput. Linguistics*, Santa Fe, NM, USA, 2018, pp. 3715–3727.

[3] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1–9.

[4] Z. Liu, W. Yu, W. Chen, S. Wang, and F. Wu, "Short text feature selection for micro-blog mining," in *Proc. Int. Conf. Intell. Softw. Eng.*, Wuhan, China, Dec. 2010, pp. 1–4.

[5] C.-D. Zhang, X. Wu, M.-L. Shyu, and Q. Peng, "A novel Web video event mining framework with the integration of correlation and co-occurrence information," *Comput. Sci. Technol.*, vol. 28, no. 9, pp. 788–796, Sep. 2013.

[6] C. Zhang, X. Wu, M.-L. Shyu, and Q. Peng, "Integration of visual temporal information and textual distribution information for news Web video event mining," *IEEE Trans. Human-Mach. Syst.*, vol. 46, no. 1, pp. 124–135, Feb. 2016.

[7] C. Zhang, D. Liu, X. Wu, G. Zhao, M.-L. Shyu, and Q. Peng, "Near-duplicate segments based news Web video event mining," *Signal Process.*, vol. 120, pp. 26–35, Mar. 2016.

[8] X. Yan, J. Guo, Y. Lan, and X. Cheng, "A biterm topic model for short texts," in *Proc. 22nd Int. Conf. World Wide Web*, Rio de Janeiro, Brazil, 2013, pp. 1445–1456.

[9] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1915–1929, Aug. 2013.

[10] M. Alhawarat and M. Hegazi, "Revisiting K-means and topic modeling, a comparison study to cluster arabic documents," *IEEE Access*, vol. 6, pp. 42740–42749, 2018.

[11] F. Wang, K. Orton, P. Wagenseller, and K. Xu, "Towards understanding community interests with topic modeling," *IEEE Access*, vol. 6, pp. 24660–24668, 2018.

[12] G. Xu, Z. Yu, and Q. Qi, "Efficient sensitive information classification and topic tracking based on tibetan Web pages," *IEEE Access*, vol. 6, pp. 55643–55652, 2018.

[13] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Commun. ACM*, vol. 18, no. 11, pp. 613–620, 1975.

[14] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *J. Amer. Soc. Inf. Sci.*, vol. 41, no. 6, pp. 391–407, 1990.

[15] T. Hofmann, "Probabilistic latent semantic indexing," in *Proc. 22nd Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Berkeley, CA, USA, 1999, pp. 50–57.

[16] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.

[17] X. Guo, Y. Xiang, Q. Chen, Z. Huang, and Y. Hao, "LDA-based online topic detection using tensor factorization," *Inf. Sci.*, vol. 39, no. 4, pp. 459–469, Mar. 2013.

[18] D. Ramage, D. Hall, R. Nallapati, and C. D. Manning, "Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Singapore, 2009, pp. 248–256.

[19] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, "Hierarchical Dirichlet processes," *J. Amer. Statist. Assoc.*, vol. 101, no. 476, pp. 1566–1581, Dec. 2006.

[20] T. Chou and S. Cheng, "Design and implementation of a semantic image classification and retrieval of organizational memory information systems using analytical hierarchy process," *Omega*, vol. 34, no. 2, pp. 125–134, 2006.

[21] S.-C. Cheng, T.-C. Chou, C.-L. Yang, and H.-Y. Chang, "A semantic learning for content-based image retrieval using analytical hierarchy process," *Expert Syst. Appl.*, vol. 28, no. 3, pp. 495–505, Apr. 2005.

[22] J. Fan, Y. Gao, H. Luo, and G. Xu, "Statistical modeling and conceptualization of natural images," *Pattern Recognit.*, vol. 38, no. 6, pp. 865–885, Jun. 2005.

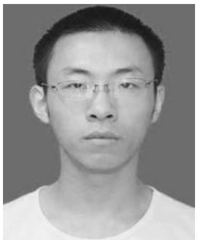
[23] J. Luo, A. E. Savakis, and A. Singhal, "A Bayesian network-based framework for semantic image understanding," *Pattern Recognit.*, vol. 38, no. 6, pp. 919–934, Jun. 2005.

[24] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006, doi: 10.1126/science.1127647.

[25] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, Jun. 2009, pp. 248–255.

CHENGDE ZHANG received the M.S. degree in computer application and technology from Xihua University, Chengdu, China, in 2009, and the Ph.D. degree in Southwest Jiaotong University, Chengdu, China, in 2015. From 2012 to 2013, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Miami, USA. He is currently a Lecturer with the School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan, China. His research interests include multimedia information retrieval, data mining, image processing, and pattern recognition.

SHAOZHEN LU received the B.Eng. degree from the School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan, China, in 2016. His research interests include topic detection, data mining, deep learning image processing, and image understanding.



CHENGMING ZHANG received the B.Eng. degree from the School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan, China, in 2016. His research interests include image processing, data mining, and topic detection.

XIA XIAO received the master's degree from Bubei University, Wuhan, China, in 2013. From 2013 to 2018, she was a Lab Assistant with the Wuhan Institute of Virology, CAS, Wuhan. She is currently a Lab Assistant with the Laboratory Management Center, Wuhan Qingchuan University, Wuhan. Her current research interests include image processing and data mining.

QIAN WANG received the B.S. degree from the Department of Computer Science and Technology, Central China Normal University, Wuhan, China, in 2005, and the Ph.D. degree from the School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, in 2009. She is currently an Associate Professor with the School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan. Her current research interests include computer vision, computer aided diagnosis, and pattern recognition. She has published about 15 academic papers in related journals and conferences. She also chairs several research projects, including a project of the National Natural Science Foundation of China.

GAO CHEN, received the masterŠs degree from Xiamen University, Xiamen, China, in 2009, and the Ph.D. degree from Southwest Jiaotong University, Chengdu, China, in 2016. From 2016 to 2018, he was a Postdoctoral Researcher with the Department of Electronic Engineering, Tsinghua University, Beijing, China. He is currently a Lecturer with the School of Electrical Engineering and Intelligentization, Dongguan University of Technology, Dongguan, China. His current research interests include remote sensing image processing and multimedia image processing.