Novel Visual and Statistical Image Features for Microblogs News Verification

Zhiwei Jin, Juan Cao, Yongdong Zhang, *Senior Member, IEEE*, Jianshe Zhou, and Qi Tian, *Fellow, IEEE*

***Abstract*—Microblog has been a popular media platform for reporting and propagating news. However, fake news spreading on microblogs would severely jeopardize its public credibility. To identify the truthfulness of news on microblogs, images are very crucial content. In this paper, we explore the key role of image content in the task of automatic news verification on microblogs. Existing approaches to news verification depend on features extracted mainly from the text content of news tweets, while image features for news verification are often ignored. According to our study, however, images are very popular and have a great influence on microblogs news propagation. In addition, fake and real news events have different image distribution patterns. Therefore, we propose several visual and statistical features to characterize these patterns visually and statistically for detecting fake news. Experiments on a real-world multimedia dataset collected from Sina Weibo validate the effectiveness of our proposed image features. The news verification performance of our method outperforms baseline methods. To the best of our knowledge, this is the first attempt that systematically explores image features on news verification task.**

***Index Terms*—Fake news detection, image features, microblogs, news verification, rumor detection, social media.**

I. Introduction

As ONE of the most popular social media services, microblog, such as Twitter and Chinese Sina Weibo, provides first-hand news posted by traditional journalists and common users. Users on microblogs not only consume news but also propagate and even produce immediate news on the social network. With millions of people serving as "news sensors", news on Microblog is valuable for opinion mining and decision making.

However, the convenience of publishing news also fosters the emergence of various fake news. Without verifying the truthfulness of news, fake news would spread promptly through social network and result in serious consequences [1]. According to a report in China, more than 1/3 trending events on microblogs contain fake information [2].

To detect fake news on microblogs, existing studies extract features from various aspects of news tweets. With these features, classification-based methods [3]–[8] or graph-based optimization methods [9], [10], [11] are deployed to identify news as real or fake. The extraction of specific and effective features is quite important for news verification.

Considering the characteristics of news in microblogs, various features are proposed in the existing literature. Generally, these features are extracted from tweets text content [3], tweets propagation process [5], [7] and involved users [3], [4], *etc*. However, image features, as a very important set of features, are ignored by them. We find it necessary to incorporate image features for news verification for two reasons:

On one hand, images are very popular in microblogs. As a multimedia social network, microblogs are full of images. Constrained by the text length limitation (140 characters or less per tweet), it is more efficient to tell a news story with attached images. In fact, in our real-world data set collected from Sina Weibo (Table I), the ratio of image to tweet is more than 0.516. This indicates more than half of tweets come along with images on average. Recently, many studies explore these multimedia contents on social media for understanding social events, such as multi-modal topic model [12], multimedia summarization on microblogs [13], election prediction [14], popularity prediction [15], visual concept learning [16] and opinion mining [17]. For understanding the truthfulness of social news events, it is important to exploit visual contents on social media.

On the other hand, images have considerable impact on news diffusing in microblogs. Images reveal vivid descriptions of the news event situations so that they attract more attentions than pure text content. We define the tweet with image attached as image-tweet and the pure text tweet as text-tweet. In our proposed data set, the average re-tweet number of image-tweet is more than 11 times larger than that of text-tweet (191 v.s. 16). This difference reflects that images play an important role in the news diffusion process.

In addition to their popularity and importance, images also have different distribution patterns for fake and real news on

Manuscript received February 29, 2016; revised June 2, 2016 and July 13, 2016; accepted October 3, 2016. Date of publication October 12, 2016; date of current version February 14, 2017. This work was supported in part by the National Nature Science Foundation of China under Grant 61525206, Grant 61571424, and Grant 61428207, and in part by the Beijing Advanced Innovation Center for Imaging Technology under Grant BAICIT-2016009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Benoit Huet.

Z. Jin and J. Cao are with the Key Laboratory of Intelligent Information Processing of the Chinese Academy of Sciences, Institute of Computing Technology, Beijing 100190, China (e-mail: jinzhiwei@ict.ac.cn; caojuan@ict.ac.cn).

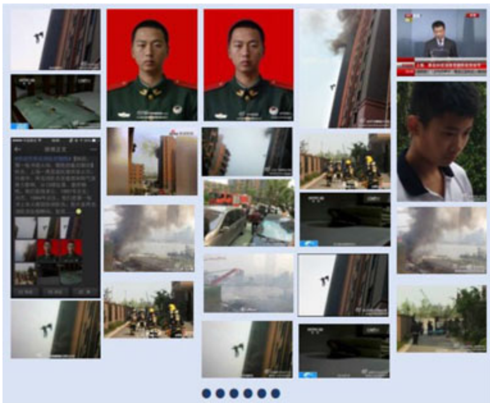
Y. Zhang is with the Key Laboratory of Intelligent Information Processing of the Chinese Academy of Sciences, Institute of Computing Technology, Beijing 100190, China, and also with the Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing 100048, China (e-mail: zhyd@ict.ac.cn).

J. Zhou is with the School of Literature, Capital Normal University, Beijing 100089, China (e-mail: zhoujs@mail.cnu.edu.cn).

Q. Tian is with the Department of Computer Science, University of Texas at San Antonio, San Antonio, TX 78249 USA (e-mail: qitian@cs.utsa.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2016.2617078



(a)

(b)

Fig. 1. Images in the real and fake news examples. It is obvious that images in the real news are much more diverse than those in the fake news. Such difference of image distribution patterns in news events can be utilized to distinguish real and fake news. (a) Images in the real news example. (b) Images in the fake news example.

microblogs. Intuitively, people tend to report news with images taken by themselves at the event scene. If the event is real, then various images taken by different witnesses would be posted on microblogs. If the event is fake, there are very few images or repeatedly posted images.

Images in fake and real news are visually distinctive. In Fig. 1, we illustrate images in two pieces of news as examples: one is a real news event[1] and the other is a fake news event.[2] Although both news events are emergency events: the real one is about a fire accident and the fake one is about a flight accident, image distributions of them are quite different. It is obvious that images in the real news example are more diverse than those in the fake news.

Meanwhile, images also have statistically distinctive patterns for fake and real news. For each news event in our data set, we plot the number of tweets and the number of images it contains (Fig. 2). We can observe that the image-to-tweet ratio of real news tends to be larger than that of fake news: given the same number of tweets in news events, real news events trend to contain more images than fake news events. In other words, images in the real news are much denser than those in the fake news.

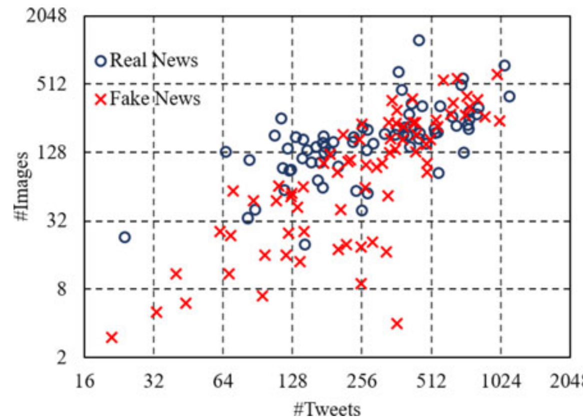


Fig. 2. Relationship of the number of tweets and the number of images in fake and real news.

Based on above observations, we assume that images have important impact on detecting fake news in microblogs. Thus, we propose novel image features for news verification on microblogs in this paper. Image features are extracted from two aspects: the visual characteristics and the overall statistics of images in news event. To exploit visual content of images, visual features are extracted to describe image distributions from a visual perspective, such as visual clarity, diversity and coherence features in a news event. The statistical features capture basic statistics of images, such as image number, multi-image number, image-to-tweet ratio, *etc*. Some attributes of images, such as resolution and popularity, are also extracted as statistical features.

In summary, we focus on image features for news verification in this work. To the best of our knowledge, this is the first attempt that systematically explores image features to verify news on microblogs. The main contributions are summarized as follows:

1) We propose a set of visual features for news verification on microblogs. Extracted from image visual content, these features reveal hidden characteristics of image distributions in news events and proved to be very effective for detecting fake news.
2) We propose several image statistical features as complements to visual features. These statistical features summarize image statistics and attribute information in news event. They capture image distribution patterns quantitatively and can further improve the news verification performance.
3) We collect a real-word multimedia data set from Sina Weibo for performance evaluation. In this data set, 50,287 tweets and 25,953 images in fake and real news events on Sina Weibo are collected. The data set is comparable in size with most of the recently released data sets and contains ground truth from authoritative sources for a fair evaluation.

The remainder of this paper is organized as follows. In Section II, we give an overview of related work. In Section III, formal definition of the problem is presented. A real-word multimedia data set, upon which analysis and experiments are done,

[1]Two fire fighters died in battling a fire in Shanghai. [Online]. Available: http://weibo.com/2286908003/B2inhn2bl

[2]The U.S. military deployment in Thailand listened an SOS signal from the missing Flight MH370. [Online]. Available: http://weibo.com/3261552423/AA4cExgrZ

is also proposed in this section. We give details of proposed visual features in Section IV, followed by statistical features in Section V. Details of data set and experimental results are presented in Section VI. We conclude the paper in Section VII.

II. Related Work

Given our focus on effective features for news verification on microblogs, we present a brief review of related work from two research streams: non-image features for news verification and image features for news verification.

A. Non-image Features for News Verification

Existing studies present a wide range of features for detecting fake news on microblogs. We summarize these features in three categories: text content features, user features and propagation features.

Text content features are features extracted from tweet text statistically or semantically. Statistical features capture basic statistics of tweets, such as tweet length, word count, and punctuation, containing of URLs or hash tag topics [3] and types of emoticons used and POS tags [18]. Semantic features include sentiment scores [3], [19] of tweets, and opinion words [5] in tweet text. Previous researches [3], [5] show that not all these features are effective for news verification. Semantic features are highly dependent on the performance of text semantic mining, which is not an easy task itself.

User features are extracted from tweet publishers, such as the time and location of the account registration, gender and age of the user [20], the verification type of the account, number of friends, number of followers, the description and the personal home page of the user, number of messages post in the past [3]. Although user features are crucial in this task, they may be unreliable sometimes. In [10], the authors argue that even authorized users participate in the propagation of many fake news events.

Propagation features are statistics of the propagation process of news events, such as the total number of nodes in the propagation graph, the maximum and average degree of the root node [3], [5], [21]. Wu *et al.* [7] proposed a concise structure to describe the propagation process for a single tweet. However, their work is on the messages level. How to incorporate it in the event level news verification has not been explored yet.

Apart from above features, some other features, such as type of software client, location from which a tweet is posted [4] are also presented in literature. In practice, features from various types are fused to give a stable verification result [22].

B. Image Features for News Verification

Images are very popular on microblogs, but how they are related to the credibility of tweets are not sufficiently explored. Morris *et al.* [20] release a survey result that user profile image has important impact on information credibility published by this user. Similarly, Gupta *et al.* [9] define a user feature to record whether the user has a profile image for evaluating the credibility of user.

For images attached in tweets, only very basic features are proposed in literature. A tweet-level feature "has multimedia" is defined in [7] to record the status of multimedia attachment of a tweet: whether the tweet has any picture, video or audio attached. Gupta *et al.* [23] make an effort to understand the temporal, social reputation and influence patterns for the spreading of fake images on mircoblogs. They propose a classification model to identify the fake images on Twitter during Hurricane Sandy. Some interesting conclusions are drawn: The original fake images are limited, and 86% of fake images were re-tweets. These conclusions meet our assumption that images in the fake news are less diverse and limited in amount. However, their work is still based on traditional text, user, and propagation features. In [6], the authors point out that the fake news is more likely to contain outdated images published earlier. They propose a time span feature to capture this time delay. The Baidu search engine is deployed to find the original image for calculating this time span. According to their results, this feature is quite effective. However, only a small part of the fake news containing outdated images on microblogs.

Aiming to automatically predict whether a tweet that shares multimedia content is fake or trustworthy real, Boididou *et al.* [24] proposed the Verifying Multimedia Use task that took place as part of the 2015 MediaEval benchmark. Several user, content and image forensics [25] features are used as baseline features for this task. This task attracts attentions for verifying images in tweets. As a solution to this task, the work of Jin *et al.* [8] mostly focused on classification model for the problem rather than image features.

It is obvious that none of above work explores image visual content or complicated image statistical features. To describe image distribution features visually, we resolve to the research of image quality estimation for image retrieval system. Based on their success in image retrieval, these visual features, such as clarity score [26], [27], coherence score [28], [29] and diversity score [30]–[32], are expected to be effective in news verification task.

III. Problem Definition and Data Set

A. Problem Definition

This paper studies the problem of automatic news verification on microblogs. Here news verification means to verify the truthfulness of a news event.

In the study of event detection task over broadcast news [33], the news event is commonly defined as "something that occurs in a certain place at a certain time". Considering the circumstance of microblogs, we give a formal definition for news event for our problem.

Definition of news event: A news event on microblogs is composed of a set of tweets containing certain keywords during a certain period of time.

With this definition, the keywords specify the news content like what it is about, where it happens; and the time scope specifies when the event happened. For example, considering the real news example in Section I "*Two fire fighters died in battling a fire in Shanghai*", we obtain its keywords as "fire fighter, Shanghai" and its time span as "2014-04-3 – 2014-05-02".

Fig. 3. Example of a tweet on Sina Weibo with some of its key features marked: user, text content, social context (re-tweets, comments, and likes), and image content (two images).

Tweets which are containing these keywords and posted during this time period constitute this news event.

Another important concept in this definition is "tweet". A tweet on microblogs is a short message posted by a microblog user. This message can contain hash tag topics (#), external URL links or reference to other users (@). Besides, users can also post one or more images along with their messages. Once the tweet is released, it may draw attention over the social network and be re-tweeted(re-posted), commented or liked by other users. Re-tweets, comments and likes form a social context for the target tweet. Thus, we define a tweet as follows:

Definition of tweet: A tweet on microblogs is a piece of message posted by a certain user along with social context and images (if it contains any).

In this definition, a tweet is specifically defined with its features from several aspects: text message, user, social context and image content. The former three types of features are extensively studied in existing literature, and little work has been done on the effect of images for news verification. Fig. 3 illustrates a tweet from Sina Weibo with its several key features marked.

B. Data Set

Credibility verification of multimedia content on social media is a fairly new problem. Very few data sets are publicly available. As the most similar work to this paper, the Verifying Multimedia Use task at MediaEval 2015 [24] released a dataset collected from Twitter. This set is designed for message-level verification. It contains only 191 distinct images from 17 events. And only 5 of them contain more than 10 images. However, we aim to verify news at the event-level and are especially interested in the distribution patterns of images in news events, such as diversity or coherence patterns. Without sufficient images in the event, it could lead to very biased results.

In order to detect image patterns in depth, we collect a multimedia data set with sufficient images from diverse news events (Fig. 4). Compared with previous work, this data set is also featured in its objective ground truth. While most existing studies present the evaluation data set with ground truth defined by human annotators [3], [5], we use the decision of authoritative sources to form a convincing ground truth of the truth value of news events in the data set. Specifically, for the collection of fake

(a)

(b)

Fig. 4. Images in various news events of proposed data set. Images are resized to the same size for display. (a) Images in real news events. (b) Images in fake news events.

TABLE I
DETAILS OF PROPOSED DATA SET

	Fake News	Real News	All
Count	73	73	146
#Tweets	23 456	26 257	49 713
#Images	10 231	15 287	25 513
#Distinct User	21 136	22 584	42 310

news events, the official rumor busting system of Sina Weibo[3] is the main source [7], [10]. Each news on the system is carefully examined and the ones that are not fake or not news-worthy are deleted by the human expert; for the collection of real news events, a hot news detection system of Xinhua News Agency, the official and most authoritative news agency in China, is used as the source.[4]

For each news event, its keywords and time span are extracted. Then a query is formed with specified keywords and time span. Tweets returned by Weibo search engine[5] answering this query are collected. Images, user information and social context information of all tweets are crawled accordingly. After removing duplicated news, news with fewer than 20 tweets and news without any images, we get 73 fake news events. To keep a balanced data set, we randomly choose 73 real news events from over 7,000 crawled real news events as negative samples in the data set.

Finally, our proposed data set is constituted of 146 news events, 50,287 tweets (19,762 of them have images attached), 25,953 images and 42,441 distinct users (Table I). Apart from its multimedia content and fair ground truth, this data set is comparable in size to those in existing studies in terms of news

[3][Online]. Available: http://service.account.weibo.com/

[4]This system is not publicly available yet at the moment. Equivalent replacement can be other authoritative news sources, such as hot news recommendation system of Sina Weibo.

[5][Online]. Available: http://s.weibo.com/

TABLE II
DESCRIPTIONS OF FIVE VISUAL FEATURES

Feature	Description
Visual Clarity Score	The distribution difference between event image set and collection image set.
Visual Coherence Score	The average of visual similarities between image pairs in event image set.
Visual Similarity Distribution Histogram	The distribution histogram of the image visual similarity matrix.
Visual Diversity Score	The weighted average of visual dissimilarities between image pairs.
Visual Clustering Score	The number of image clusters after cluster images with visual patterns.

events while contains nearly 10 times more tweets. For example, in most recent studies, the two data sets in [9] have 67 and 83 fake events respectively and the results in [7] are reported on a set of about 5000 tweets.

IV. Visual Features for News Verification

Images have discriminative visual distribution patterns fake and real news events, but how to describe image distributions based on their visual content? In the study of image retrieval, some visual features are proposed [26]–[28], [30], [34] to estimate the distribution of images in the search results. The success of these features for estimating image distributions inspires us to construct similar features for quantitatively describing image distribution patterns of images in microblog news.

Specifically, we propose five visual features (Table II): visual clarity score, visual coherence score, visual similarity distribution histogram, visual diversity score, and visual clustering score. These features describe image distribution characteristics from different visual aspects and reveal hidden distribution patterns of images in news events.

A. *Visual Clarity Score*

The visual clarity score [27] measures the distribution difference between two image sets: one is the image set in a certain news event (event set) and the other is the image set containing images of all events (collection set). The logic behind this feature is simple: if an event set is distant from the collection set, then this event probably contains general images and is likely to be a real news event.

To compute this score, two language models are calculated: one is the language model derived from target event set and the other is the language model derived from collection set containing images of all events. Then clarity score is defined as the Kullback-Leibler divergence [35] between two language models. Kullback-Leibler divergence(KL divergence) is a commonly used measure of the difference between two probability distributions.

We use the bag-of-word image representation to define language models for images. Firstly, local descriptors for each image are extracted. Secondly, all descriptors are quantized to form a visual word vocabulary. Finally, each image is represented by words from the vocabulary. In this paper, the scale-invariant feature transform (SIFT) [36] feature are used as the local descriptor. We then use an efficient vocabulary tree building method [37] to quantize each SIFT descriptor into its corresponding visual word to form a visual word vocabulary V. Finally, we get a sequence of visual words to represent each image. Based on this representation, we estimate the event language model $P(w|k)$ and collection language model $P(w|c)$ as follows.

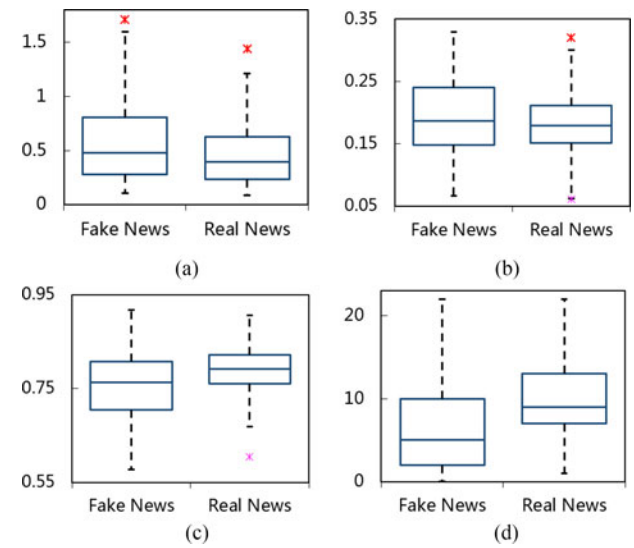


Fig. 5. Box plots for visual features of fake and real news in proposed data set: (a) visual clarity score, (b) visual coherence score, (c) visual diversity score, and (d) visual clustering score.

The news language model is defined as

$$P(w|k) = \sum_{x \in R} P(w|x)\, P(x|k). \tag{1}$$

Here, $w \in V$ is a visual word, and R is the image set of the target event specified by k (the news keywords and time scope). $P(w|x)$ is defined as the term frequency of the visual word w in the image x. For $P(x|k)$

$$P(x|k) \propto P(k|x)\, P(x). \tag{2}$$

We assume each image x has an equal prior $P(x)$, so we only need to estimate the likelihood $P(k|x)$. As k is the identifier of the news event and each image can belong to only one news event, we define $P(k|x) = 1$ if image x appears in the image set R of the news specified by k, else $P(k|x) = 0$

$$P(k|x) = \begin{cases} 1, & \text{if } x \in R \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

For the language model of collection set, $P(w|c)$ is defined as the term frequency of visual word w over images in all news events. Then visual clarity score is defined as the KL divergence between the news language model and collection language model

$$\begin{aligned} \text{VCS} &= D_{KL}\left(P(w|k)\,||\,P(w|c)\right) \\ &= \sum_{w \in V} P(w|k) \log \frac{P(w|k)}{P(w|c)}. \end{aligned} \tag{4}$$

To inspect the discriminative capacity of visual features and give some insight to how they work, we draw box plots (Fig. 5) for them based on real and fake news events in our proposed data set. As Fig. 5(a) shows, visual clarity score reveals the difference between two classes: the real news events tend to have lower visual clarity score than the fake news events. Visual clarity score is defined as the distribution difference of images between the target news event and all events. Thus, the bigger this score is, the more distinct the target news is. By assuming that images in the real news events are from various sources and images in the fake news events are from limited sources, it is comprehensible that images in real news events are more general and have lower visual clarity score. Therefore, the difference of VCS in two classes revealed in Fig. 5(a) would be useful for separating real and fake news events.

B. Visual Coherence Score

The visual coherence score[28], [29] measures how coherent images in a certain news event are. Since relevant images share common visual patterns, they are more visually similar than irrelevant images. By calculating the visual coherence score, we can measure similarity features of images in news events quantitatively.

To compute visual coherence score, we examine the visual similarity between any image pair within images in the target event image set.

$$\text{VCoS} = \frac{1}{|N(N-1)|} \sum_{i,j=1,\ldots,T; i \neq j} sim(x_i, x_j) \quad (5)$$

Here N is number of the images in R, $sim(x_i, x_j)$ is the visual similarity between image x_i and image x_j. In this paper, we compute the similarity between images based on their GIST features [38].

GIST is a popular global descriptor for image and finds its successful application in image retrieval and scene recognition. To represent the dominant spatial structure (naturalness, openness, roughness, expansion, ruggedness) of a scene, a set of perceptual dimensions are proposed for this feature. To extract GIST feature for an image, the image is divided into a 4-by-4 grid for which orientation histograms are extracted. We use the implementation of [38] to compute a 512-dimension GIST feature vector for each image and use the cosine similarity metric to compute image similarity $sim(x_i, x_j)$.

As the box plot of visual coherence score shows [Fig. 5(b)], the fake news events tend to have slightly higher visual clarity score than the real news events. This means images in fake news events are more coherent and less diverse, which also meets the diversity difference assumption of fake and real news events. Consequently, this feature can be useful for detecting fake news.

C. Visual Similarity Distribution Histogram

While the visual coherence score evaluates the overall similarity characteristics of images in news event, we propose a fine granularity feature as complementary. Compared with visual coherence score, this feature describes image similarity distribution with more detailed computations and represents it with a set of values.

This feature is computed based on the whole similarity matrix of all images in a target news event [29].

For a given set of news images R, a visual similarity matrix $\mathbf{M} \in \mathbb{R}^{N \times N}$ is obtained by calculating pairwise image similarity. Each $m_{i,j} \in \mathbf{M}$ denotes the visual similarity between the i-th ranked image and the j-th ranked image in R. Images in a news event is ranked according to their popularity: images with more re-tweets and comments are ranked higher. This comes from the assumption that images arousing a lot of social attentions are representative images in a news event.

The visual similarity $m_{i,j}$ is computed the same as (5) and it is in the range of $[0, 1]$. Then, the image similarity matrix is quantified into an H-bin histogram by mapping each element in the matrix into its corresponding bin. The number of bins is set as 10 in this paper. This visual similarity distribution histogram is denoted as VSDH

$$\text{VSDH}(h) = \frac{1}{N^2} |\{(i,j)\,|i,j \leq N, m_{i,j} \in h\text{-}th\ bin\}|, h = 1, \ldots, H. \quad (6)$$

As a detailed version of visual coherence score, VSDH is assumed to have positive effect for news verification.

D. Visual Diversity Score

The visual diversity score measures the visual difference in the image set of a target news event. Compared with visual coherence score, it computes the image diversity distribution directly and gives more emphasis on representative images.

In image search task, the diversity of an image is defined as its minimal difference with the images ranking before it [31]. Similarly, assuming a ranking of images $x_1, x_2, \ldots, x_N$ in the event image set R, the diversity score [30] of all images in R is defined as

$$\text{VDS} = \sum_{i=1}^{N} \frac{1}{i} \sum_{j=1}^{i} dis(x_i, x_j). \quad (7)$$

Here, $dis(x_i, x_j)$ denotes the dissimilarity between images ranked in the position i and j. We define it as the complementary of similarity

$$dis(x_i, x_j) = 1 - sim(x_i, x_j). \quad (8)$$

Here, $sim(x_i, x_j)$ has the same definition as in (5). Images are also firstly ranked by their popularity based on the assumption that popular images have better representation for the news event.

While visual coherence score is the average similarity over all images, the visual diversity score is a weighted average of dissimilarity over all images. In fact, the diversity score gives more emphasis on top-ranked images. So that VDS can reduce the impact of noise images. We can observe in the box plot for this feature [Fig. 5(c)] that the real news events tend to have higher visual diversity score than the fake news events. Comparing Fig. 5(b) and 5(c), we find that VDS is a even better feature than VCoS for separating fake and real news events.

E. Visual Clustering Score

The visual clustering score (VClS) evaluates the distribution of images in news event from a clustering perspective. It is

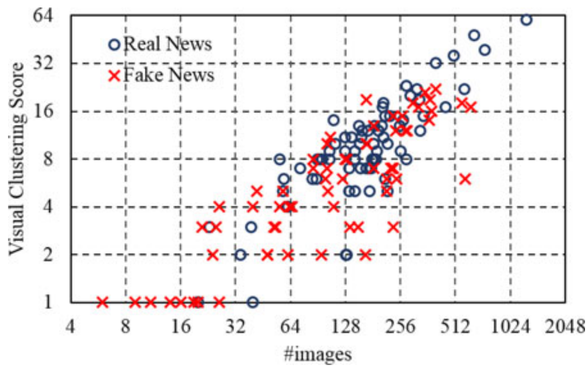


Fig. 6. Relationship of the number of images and visual clustering score in fake and real news.

defined as the number of clusters formed by all images in a target news event.

To cluster images, the hierarchical agglomerative clustering (HAC) algorithm [39] is employed. This algorithm forms clusters with a bottom-up strategy by merging nearest atomic clusters into larger clusters. The single-link strategy is used to measure the similarity between two clusters, which is defined as the nearest distance of object pairs in them. We use the Euclidean distance of image GIST feature vectors as the distance measurement. Unlike some other clustering algorithms, such as K-means [40], HAC does not require the number of clusters to be determined or estimated before clustering. Under the same threshold of link distance for all events, the number of clusters yielded by HAC reveals the diversity feature of images. After clustering, very small clusters, i.e. clusters with no more than 3 images, are considered as outliers and removed. Then we define the visual clustering score as the number of remained clusters.

As the box plot of visual clustering score [Fig. 5(d)], images in the fake news events tend to form fewer clusters than those in the real news events. This observation supports the image diversity difference from the clustering angle and indicates that this feature can be useful for separating fake and real news events. In addition, we plot the relation of visual clustering score of a news event and the number of images it contains (Fig. 6) to argue that fake news event have lower clustering score not purely because they have fewer images: given the same number of images in a news event, fake news events also trend to form fewer image clusters than real news events. In other words, image in fake news events are less diverse than those in real news events.

V. Statistical Features for News Verification

In existing studies [3], statistical features of text content are extracted for detecting fake news on microblogs and some of them are proved to be useful. Therefore, we assume similar statistical features of image content can be also effective in this task. In fact, as Fig. 2 shows, the basic image statistics, "image-tweet ratio", is a distinctive feature for distinguish real and fake news.

TABLE III
Descriptions of Seven Statistical Features

Feature	Description
Count	The number of all images in a news event.
Image Ratio	The ratio of the image-tweets in all tweets.
Image Ratio II	The ratio of image number to tweet number.
Multi-image Ratio	The ratio of multi-image-tweets in all tweets.
Multi-image Ratio II	The ratio of multi-image-tweets in all image-tweets.
Hot Image Ratio	The ratio of the most popular image in all distinct images.
Long Image Ratio	The ratio of tweets with long image in all image-tweets.

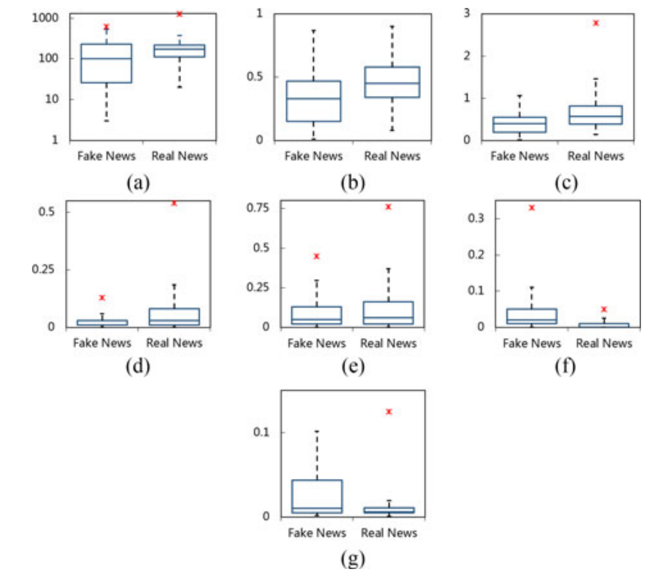


Fig. 7. Box plots for seven statistical features of fake and real news in proposed data set: (a) count, (b) image ratio, (c) image ratio II, (d) multi-image ratio, (e) multi-image ratio II, (f) hot image ratio, and (g) long image ratio.

We give some inspection of images posted on microblogs: 1) users can post zero, one or more than one images along with a text content in a tweet; 2) some images are very popular and gain a lot of re-tweets and comments; 3) most images are in regular resolutions, while a few images are not. Based on these observations, we define different types of tweets and images from three aspects:

1) *Number of images in a tweet:* Text-tweet: tweet without image attached; image-tweet: tweet with at least one image; multi-image-tweet: tweet with more than one image.
2) *Popularity:* Hot image: the most popular image in a news event. Popularity is defined as the number of re-tweets and comments gained by this image-tweet.
3) *Long image:* image with a length-to-width ratio larger than 1.9. It is often a collage of several images.

Basic statistics of these types are extracted as statistical features for images in news events. As a result, we get a total of seven statistical features for describing image distribution characteristics (Table III). From the box plots of these features (Fig. 7), we observe that all these features have discriminative distributions on real and fake news to some degree. Some of them are prominent, such as "count" [Fig. 7(a)] and "image ratio" [Fig. 7(b)]. However, compared with visual features [Fig. 5], some statistical features are inclined to be affected by outliers,

TABLE IV
42 EXISTING FEATURES

Type	Feature
Text Content	Count of Message, Average Word/Character Length, Fraction of Question/Exclamation Mark, Fraction of Multi Question/Exclamation Mark Ratio, Fraction of First/Second/Third Pronouns, Fraction of URL/@/#, Count of Distinct URL/@/#, Fraction of Popular URL/@/#, Count of Distinct People/Location/Organization, Fraction of People/Location/Organization, Fraction of Popular People/Location/Organization, Average Sentiment Score, Fraction of Positive/Negative Tweets.
User	Count of Distinct Users, Fraction of Popular Users, Average Followers/Followees/Posted Tweets, Fraction of Verified User/Organization.
Propagation	Size of Max Subtree, Average Likes, Average Degree/Non-zero Degree.

TABLE V
PERFORMANCE COMPARISONS ON OUR PROPOSED DATA SET

Model	Feature Set	Accuracy	Fake News			Real News		
			Precision	Recall	F_1	Precision	Recall	F_1
SVM	Castillo [3]	0.651	0.657	0.63	0.643	0.645	0.671	0.658
	Kwon [5]	0.699	0.704	0.685	0.694	0.693	0.712	0.703
	NonImg	0.719	0.695	**0.781**	0.735	0.75	0.658	0.701
	NonImg + Img	**0.795**	**0.812**	0.767	**0.789**	**0.779**	**0.822**	**0.8**
Logistic Regression	Castillo	0.589	0.603	0.521	0.559	0.578	0.658	0.615
	Kwon	0.719	0.722	0.712	0.717	0.716	0.726	0.721
	NonImg	0.753	0.747	**0.767**	0.757	0.761	0.74	0.75
	NonImg + Img	**0.781**	**0.789**	**0.767**	**0.778**	**0.773**	**0.795**	**0.784**
KStar	Castillo	0.637	0.635	0.644	0.639	0.639	0.63	0.634
	Kwon	0.671	0.647	0.753	0.696	0.705	0.589	0.642
	NonImg	0.658	0.635	**0.74**	0.684	0.689	0.575	0.627
	NonImg + Img	**0.726**	**0.732**	0.712	**0.722**	**0.72**	**0.74**	**0.73**
Random Forest	Castillo	0.692	0.684	0.712	0.698	0.7	0.671	0.685
	Kwon	0.753	0.776	0.712	0.743	0.734	0.795	0.763
	NonImg	0.767	0.791	0.726	0.757	0.747	0.808	0.776
	NonImg + Img	**0.836**	**0.855**	**0.808**	**0.831**	**0.818**	**0.863**	**0.84**

such as "multi-image ratio" [Fig. 7(d)] and "long image ratio" [Fig. 7(g)].

In short, we propose five visual features and seven statistical features for characterizing image content in news tweets. These features are normalized and concatenated with other non-image features to represent news events. Then traditional classifiers are built based on these features to identify news events as real or fake.

VI. EXPERIMENTS

In this section, we conduct extensive experiments on our proposed data set to validate the effectiveness of proposed image features. First, we give a brief review of existing features and select the best features from these non-image features as baselines. We then describe experimental setup and evaluation metrics. Performance comparison and feature analysis are presented to give more insights into proposed image features for news verification on microblogs.

A. Baseline Features

We compare the effectiveness of our proposed image features with existing features in literature.

Castillo *et al.* [3] propose 15 features for evaluating the credibility of news events on Twitter. Following studies [5], [7], [9] propose other effective non-image features. To collect all these features, we get a total of 42 features from the aspects of text content, propagation and user. This feature set includes almost all existing features for this task and several new features (Table IV).

In order to find the best non-image feature set and remove redundant features, we perform feature selection with information gain ratio method [41] on proposed data set. The information gain ratio is a ratio of information gain to the intrinsic information. This method is commonly used for measuring the goodness of attributes in decision tree learning [42]. This procedure selects 11 features out of all 42 features with information gain ratio larger than zero. As Table V shows, most of the best non-image features are user-based features and propagation-based features.

B. Experimental Setup

1) Performance Measures: To judge the verification performance quantitatively, several performance measures are taken into consideration. Accuracy is the percentage of correctly identified fake and real news. It is a measure of the overall effectiveness of a given method. Precision and recall are computed for real and fake news respectively to examine the performance for each class. F_1 score, which is the harmonic mean of precision and recall, is also computed for both classes.

2) Feature Sets: We group features involved in the performance evaluation into different sets.

1) *Castillo:* 11 best features proposed by Castillo [3].
2) *Kwon:* 42 existing non-image features in literature [5], [7], [9] (Table IV).
3) *NonImg:* 11 best features selected from all 42 existing non-image features (Table V).
4) *Visu:* five visual features proposed in this paper (Table II).
5) *Stat:* seven statistical features proposed in this paper (Table III).
6) *Img:* five visual features and seven statistical features of images.

C. Performance Evaluation

1) Effectiveness of Proposed Image Features: To validate the effectiveness of our proposed image features, we conduct experiments with different feature sets on our proposed data set. For each feature set, four classification models are trained: SVM, Logistic Regression, KStar and Random Forest. The results of different classification models would reveal how effective each feature set is in general. We perform a 4-fold cross validation strategy for the training/validation process for each model.

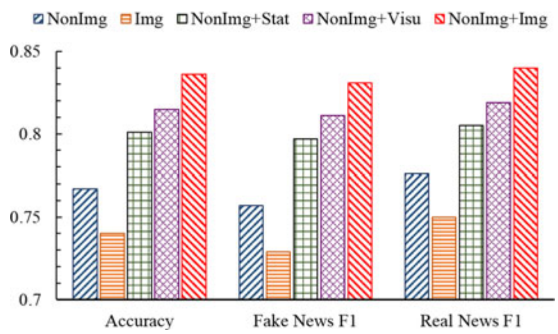


Fig. 8. Impact of visual and statistical feature sets.

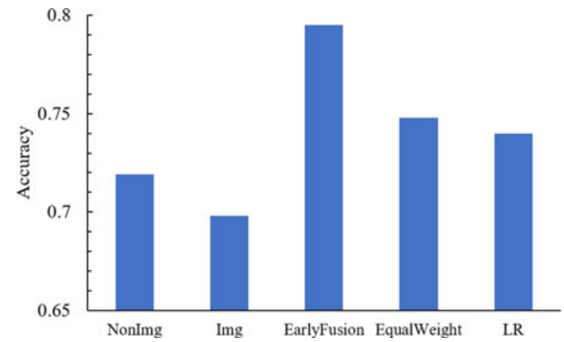


Fig. 9. Comparison of early and late fusions.

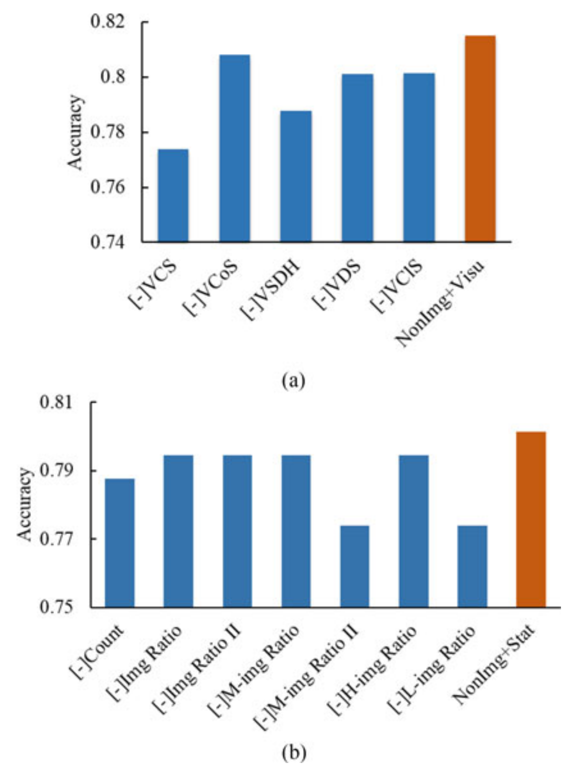


Fig. 10. Impact of individual image feature. (a) Visual features. (b) Statistical features.

From the comparison results in Table V, we can observe that:

With the image features proposed in this paper, i.e. five visual and seven statistical features of image content, all the classification models achieve much better performance compared with baseline feature sets. Among the four classification algorithms, Random Forest generates the best accuracy of 83.6%, which boosts Castillo *et al.*'s [3] by over 14% and outperforms Kwon *et al.*'s [5] by 8%. Compared with the best non-image feature set (NonImg), the accuracy of Random Forest with proposed image feature set is 7% higher, as well as fake news detection F_1 score. Similarly, the accuracy of the other three models also gains about 7% boosting with image features. These results validate the effectiveness of our proposed image features for the news verification task.

2) Impact of Visual and Statistical Feature Sets: To inspect visual and statistical feature sets separately, we experiment with Visu, Stat respectively and plot the accuracy, fake news F_1 and real news F_1 performance results in Fig. 8:

1) Our proposed statistical features (Stat) and visual features (Visu) both contribute to the performance boosting over non-image feature set. The performance increases by 3% with statistical feature set Stat and increased by 4% with visual feature set.
2) We also notice visual feature set is slightly better (1%) than statistical feature set. This meets our observation in Figs. 5 and 8 that visual features are seemly more robust than statistical features. Moreover, the impact of proposed visual and statistical feature set on the task is complementary: when the two feature sets are combined together, the performance is increased by 7%.
3) Image feature set alone results in a comparable performance to that of selected non-image feature set: the accuracy of Img is slightly inferior to that of NonImg. Considering NonImg are selected from a wide range of non-image features, this result is tolerable. Although image features play a crucial role in news verification, we should not neglect the effectiveness of non-image features.

3) Early and Late Fusion: In the above experiments, we simply concatenate proposed image features and existing non-image features into one vector. This is the so-called early fusion strategy. Late fusion is another conventional strategy for combining the representative capabilities of various features. Unlike early fusion, late fusion is based on classifier-level combination. In this experiment, we train two SVM respectively for the Img feature set and NonImg feature set. The probability prediction of two models are then combined equally weighted or further optimized through logistic regression to produce the final results.

From the comparison in Fig. 9, we can observe that the two late fusion methods, i.e., equal weight and LR, and the early fusion boost the performance of two separate feature sets in different degrees. For the two late fusion methods, the equal weight strategy is slightly better than LR. The early fusion results in the best accuracy performance among others. We find some features in the two feature sets are actually dependent. For example, the visual cluster score is correlated with total tweet number. When we train models on Img and NonImg set separately, this dependency information is missed. Therefore, the late fusion is inferior to early fusion for this problem.

D. Impact of Individual Image Feature

To study how individual image feature affects the news verification performance, we experiment with a certain feature leaving out at each time.

For each visual feature, we exclude it from the NonImg+Visu feature set and then conduct experiment with the new set following the same experimental settings as the baseline method, i.e. random forests model and 4-fold cross validation. Similarly, we conduct experiments for each statistical feature with NonImg+Stat as the baseline feature set. Then we examine the drop of accuracy with one feature excluded at each time. The degree of performance decline implies how important this feature is. Comparison results are illustrated in Fig. 10. Here, we use [-]X to denote the baseline feature set excluding feature X.

As Fig. 10(a) illustrates, each visual feature affects the performance to a different degree. Among all the visual features, VCS, VSDH, VDS and VClS have relatively large impact on the final result. Especially, without VCS, the accuracy drops from 81.5% to 77.4%. Different statistical features also affect accuracy to a different degree, as Fig. 10(b) shows. Among the seven statistical features, three of them, i.e. "count", "multi-image ratio II" and "Long image ratio", result in a decline in accuracy of more than 1% if any of them is excluded. Especially for "multi-image ratio II" and "Long image ratio", the decline in accuracy is about 3%.

VII. Conclusion

Existing features-based approaches to news verification on microblogs ignore the very important image content in tweets. In this paper, we focus on images to improve the verification performance. We find that apart from their popularity and great impact on news diffusion, images also have distinctive distribution patterns for the real and fake news visually and statistically. Therefore, image features from two aspects: visual content and statistics, are proposed to characterize the distinctiveness. Specifically, five visual features and seven statistical features are proposed. We validate the effectiveness of image features on a multimedia data set collected from Sina Weibo. Incorporating with proposed image features, our method achieves a verification accuracy of 83.6%. It significantly boosts the accuracy by more than 7% compared with baseline approaches using only non-image features.

References

[1] A. Friggeri, L. A. Adamic, D. Eckles, and J. Cheng, "Rumor cascades," in *Proc. 8th Int. AAAI Conf. Weblogs Social Media*, 2014, pp. 101–110.

[2] Z. Zhao, P. Resnick, and Q. Mei, "Enquiring minds: Early detection of rumors in social media from enquiry posts," in *Proc. 24th Int. Conf. World Wide Web*, 2015, pp. 1395–1405.

[3] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on Twitter," in *Proc. 20th Int. Conf. World Wide Web*, 2011, pp. 675–684.

[4] F. Yang, Y. Liu, X. Yu, and M. Yang, "Automatic detection of rumor on Sina Weibo," in *Proc. ACM SIGKDD Workshop Mining Data Semantics*, 2012, pp. 1–7.

[5] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang, "Prominent features of rumor propagation in online social media," in *Proc. IEEE 13th Int. Conf. Data Mining*, Dec. 2013, pp. 1103–1108.

[6] S. Sun, H. Liu, J. He, and X. Du, "Detecting event rumors on Sina Weibo automatically," in *Web Technologies and Applications*. New York, NY, USA: Springer, 2013, pp. 120–131.

[7] K. Wu, S. Yang, and K. Q. Zhu, "False rumors detection on Sina Weibo by propagation structures," in *Proc. IEEE Int. Conf. Data Eng.*, Apr. 2015, pp. 651–662.

[8] Z. Jin, J. Cao, Y. Zhang, and Z. Yongdong, "MCG-ICT at mediaEval 2015: Verifying multimedia use with a two-level classification model," in *Proc. MediaEval 2015 Multimedia Benchmark Workshop*, 2015, pp. 240–241.

[9] M. Gupta, P. Zhao, and J. Han, "Evaluating event credibility on twitter," in *Proc. SIAM Int. Conf. Data Mining*, 2012, pp. 153–164.

[10] Z. Jin, J. Cao, Y.-G. Jiang, and Y. Zhang, "News credibility evaluation on microblog with a hierarchical propagation model," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2014, pp. 230–239.

[11] Z. Jin, J. Cao, Y. Zhang, and J. Luo, "News verification by exploiting conflicting social viewpoints in microblogs," in *Proc. 13th AAAI Conf. Artif. Intell.*, Phoenix, AZ, USA, Feb. 12–17, 2016, pp. 2972–2978.

[12] S. Qian, T. Zhang, C. Xu, and J. Shao, "Multi-modal event topic model for social event analysis," *IEEE Trans. Multimedia*, vol. 18, no. 2, pp. 233–246, Feb. 2016.

[13] J. Bian, Y. Yang, H. Zhang, and T. S. Chua, "Multimedia summarization for social events in microblog stream," *IEEE Trans. Multimedia*, vol. 17, no. 2, pp. 216–228, Feb. 2015.

[14] Q. You, L. Cao, Y. Cong, X. Zhang, and J. Luo, "A multifaceted approach to social multimedia-based prediction of elections," *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2271–2280, Dec. 2015.

[15] B. Wu, T. Mei, W. Cheng, and Y. Zhang, "Unfolding temporal dynamics: Predicting social media popularity using multi-scale temporal decomposition," in *Proc. 13th AAAI Conf. Artif. Intell.*, Feb. 2016, pp. 272–278.

[16] X. Yang, T. Zhang, C. Xu, and M. S. Hossain, "Automatic visual concept learning for social event understanding," *IEEE Trans. Multimedia*, vol. 17, no. 3, pp. 346–358, Mar. 2015.

[17] Q. Fang, C. Xu, J. Sang, M. S. Hossain, and G. Muhammad, "Word-of-mouth understanding: Entity-centric multimodal aspect-opinion mining in social media," *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2281–2296, Dec. 2015.

[18] A. Hassan, V. Qazvinian, and D. Radev, "What's with the attitude?: Identifying sentences with attitude in online discussions," in *Proc. Conf. Empirical Methods Natural Language Process.*, 2010, pp. 1245–1255.

[19] V. Qazvinian, E. Rosengren, D. R. Radev, and Q. Mei, "Rumor has it: Identifying misinformation in microblogs," in *Proc. Conf. Empirical Methods Natural Language Process.*, 2011, pp. 1589–1599.

[20] M. R. Morris, S. Counts, A. Roseway, A. Hoff, and J. Schwarz, "Tweeting is believing?: Understanding microblog credibility perceptions," in *Proc. ACM Conf. Comput. Supported Cooperative Work*, 2012, pp. 441–450.

[21] M. Mendoza, B. Poblete, and C. Castillo, "Twitter under crisis: Can we trust what we RT?," in *Proc. 1st Workshop Social Media Analytics*, 2010, pp. 71–79.

[22] X. Zhou *et al.*, "Real-time news certification system on Sina Weibo," in *Proc. 24th Int. Conf. World Wide Web Companion*, 2015, pp. 983–988.

[23] A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, "Faking Sandy: Characterizing and identifying fake images on Twitter during Hurricane Sandy," in *Proc. 22nd Int. Conf. World Wide Web Companion*, 2013, pp. 729–736.

[24] C. Boididou *et al.*, "Verifying multimedia use at mediaEval 2015," in *Proc. MediaEval 2015 Workshop*, Sep. 2015, pp. 235–237.

[25] J. Li, X. Li, B. Yang, and X. Sun, "Segmentation-based image copy-move forgery detection scheme," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 3, pp. 507–518, Mar. 2015.

[26] S. Cronen-Townsend, Y. Zhou, and W. B. Croft, "Predicting query performance," in *Proc. 25th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2002, pp. 299–306.

[27] C. Hauff, V. Murdock, and R. Baeza-Yates, "Improved query difficulty prediction for the web," in *Proc. 17th ACM Conf. Inf. Knowl. Manage.*, 2008, pp. 439–448.

[28] J. He, M. Larson, and M. De Rijke, "Using coherence-based measures to predict query difficulty," in *Advances in Information Retrieval*. New York, NY, USA: Springer, 2008, pp. 689–694.

[29] X. Tian, Y. Lu, and L. Yang, "Query difficulty prediction for web image search," *IEEE Trans. Multimedia*, vol. 14, no. 4, pp. 951–962, Aug. 2012.

[30] T. Deselaers, T. Gass, P. Dreuw, and H. Ney, "Jointly optimising relevance and diversity in image retrieval," in *Proc. ACM Int. Conf. Image Video Retrieval*, 2009, Art. no. 39.

[31] M. Wang, K. Yang, X. S. Hua, and H. J. Zhang, "Towards a relevant and diverse search of social images," *IEEE Trans. Multimedia*, vol. 12, no. 8, pp. 829–842, Dec. 2010.
[32] Z. Xia, X. Wang, X. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 2, pp. 340–352, Feb. 2016.
[33] Y. Yang *et al.*, "Learning approaches for detecting and tracking news events," *IEEE Intell. Syst.*, vol. 14, no. 4, pp. 32–43, Jul./Aug. 1999.
[34] E. C. Jensen, S. M. Beitzel, D. Grossman, O. Frieder, and A. Chowdhury, "Predicting query difficulty on the web by learning visual clues," in *Proc. 28th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf.. Retrieval*, 2005, pp. 615–616.
[35] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, pp. 79–86, 1951.
[36] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
[37] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, Jun. 2006, vol. 2, pp. 2161–2168.
[38] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 145–175, 2001.
[39] B. King, "Step-wise clustering procedures," *J. Amer. Statist. Assoc.*, vol. 62, no. 317, pp. 86–101, 1967.
[40] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probability*, 1967, vol. 1, no. 14, pp. 281–297.
[41] Y. Yang and J. O. Pedersen, "A comparative study on feature selection in text categorization," in *Proc. 14th Int. Conf. Mach. Learn.*, 1997, vol. 97, pp. 412–420.
[42] S. L. Salzberg, "C4. 5: Programs for machine learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993," *Mach. Learn.*, vol. 16, no. 3, pp. 235–240, 1994.

Zhiwei Jin received the B.S. degree in software engineering from Wuhan University, Wuhan, China, in 2012, and is currently working toward the Ph.D. degree at the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, under the supervision of Prof. Y. Zhang.

His research interests include multimedia content analysis and data mining.

Juan Cao received the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2008.

She is currently working as an Associate Professor with the Institute of Computing Technology, Chinese Academy of Sciences. Her research interests include large scale social media analysis.

Yongdong Zhang (M'09–SM'13) received the Ph.D. degree in electronic engineering from Tianjin University, Tianjin, China, in 2002.

He is currently a Professor with the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. He has authored more than 100 refereed journal and conference papers. His current research interests include multimedia content analysis and understanding, multimedia content security, video encoding, and streaming media technology.

Prof. Zhang serves as an Editorial Board Member of *Multimedia Systems Journal* and *Neurocomputing*. He was the recipient of the Best Paper Award in PCM2013, ICIMCS 2013, and ICME 2010, and the Best Paper Candidate in ICME 2011.

Jianshe Zhou received the Ph.D. degree from Wuhan University, Wuhan, China, in 2002.

He is currently the Vice-President of Capital Normal University, Beijing, China, where he is also a Professor and Ph.D Supervisor. He is the Chairman of the Beijing Language Intelligence and Technology Application Collaborative Innovation Center, a Member of the International Chinese Language Society, a Reviewer of Language Subject in the Ministry of Education, and the Vice-President of the Beijing Language Society. His current academic research interests include semantics and syntactics and natural language processing.

Qi Tian (S'95–M'96–SM'03–F'16) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 1992, the M.S. degree in electrical and computer engineering from Drexel University, Philadelphia, PA, USA, in 1996, and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana Champaign, Champaign, IL, USA, in 2002.

He is currently a Professor with the Department of Computer Science, University of Texas at San Antonio, San Antonio, TX, USA. He took a one-year faculty leave at Microsoft Research Asia, Beijing, China, from 2008 to 2009. He has authored or coauthored more than 280 refereed journal and conference papers. His research projects were funded by NSF, ARO, DHS, SALSI, CIAS, and UTSA and he also received faculty research awards from Google, NEC Laboratories of America, FXPAL, Akiira Media Systems, and HP Labs. His research interests include multimedia information retrieval and computer vision.

Prof. Tian in the Editorial Board of the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON CIRCUIT AND SYSTEMS FOR VIDEO TECHNOLOGY, *Multimedia Systems Journal*, the *Journal of Multimedia,* and the *Journal of Machine Visions and Applications*. He is the Guest Editor of the IEEE TRANSACTIONS ON MULTIMEDIA, the *Journal of Computer Vision and Image Understanding*, *Pattern Recognition Letters*, the *EURASIP Journal on Advances in Signal Processing*, and the *Journal of Visual Communication and Image Representation*, among others. He was the recipient of the Best Paper Award in PCM 2013, MMM 2013, and ICIMCS 2012, and the Top 10 Candidate in PCM 2007. He was the recipient of the Research Achievement Award from the College of Science, UTSA, in 2014, and the 2010 ACM Service Award.