WILEY | Hindawi

Research Article

Visual and Textual Analysis for Image Trustworthiness Assessment within Online News

Federica Lago ⬡, Quoc-Tin Phan, and Giulia Boato

University of Trento, Department of Engineering and Computer Science, Trento 38123, Italy

Correspondence should be addressed to Federica Lago; federica.lago-1@unitn.it

The majority of news published online presents one or more images or videos, which make the news more easily consumed and therefore more attractive to huge audiences. As a consequence, news with catchy multimedia content can be spread and get viral extremely quickly. Unfortunately, the availability and sophistication of photo editing software are erasing the line between pristine and manipulated content. Given that images have the power of bias and influence the opinion and behavior of readers, the need of automatic techniques to assess the authenticity of images is straightforward. This paper aims at detecting images published within online news that have either been maliciously modified or that do not represent accurately the event the news is mentioning. The proposed approach composes image forensic algorithms for detecting image tampering, and textual analysis as a verifier of images that are misaligned to textual content. Furthermore, textual analysis can be considered as a complementary source of information supporting image forensics techniques when they falsely detect or falsely ignore image tampering due to heavy image postprocessing. The devised method is tested on three datasets. The performance on the first two shows interesting results, with F1-score generally higher than 75%. The third dataset has an exploratory intent; in fact, although showing that the methodology is not ready for completely unsupervised scenarios, it is possible to investigate possible problems and controversial cases that might arise in real-world scenarios.

1. Introduction

Images and video-audio sequences have traditionally been considered a gold standard of truth, as the process of altering or creating fake content was restricted to researchers and skilled users. With the development of tools and editing software that make the forgery process almost automatic and easy, even for nonprofessionals, this is no longer true. Not only the process of altering digital content became easier in the last years, but also the process of creating and sharing it. With more than 3 billion of users active on social media, it has been recently estimated that 3.2 billion images are shared every day, and 300 hours of video per minute are uploaded to YouTube.

In case of high-impact events, such as terrorist attacks or natural disasters, the uploaded images and videos are publicly visible and spread quickly, within seconds. The phenomenon, in which ordinary man and women are able to document events that were once the domain of professional documentary makers [1], is typically referred to as *citizen journalism*. This changes the way in which journalist and professional figures work. In fact, they no longer need to move to the location of an event, but can simply use content uploaded online. The strong competition of news outlets and individual news sources to be the first to publish news jointly with the speed of the news spreading process limits the time that journalists can spend in verifying the veracity and provenance of images and videos. The superficiality or negligence in the verification of the digital content makes the risk of spreading fake information extremely high. Moreover, fake images and videos might be shared with malicious intent [2–4] to have a higher number of clicks and thus generate a revenue (see for example [5, 6]).

This is a serious threat, as it has been proven that visual content can affect public opinion and sentiments [2, 7, 8], leading to severe consequences. For example, two spectators of the Boston Marathon were falsely portrayed as suspects of the bombings, and their picture even hit the headlines (see Figure 1(a)). This caused them emotional distress, invasion of privacy, and the risk of losing their jobs [9]. Also fake images

(a) Miscontextualized image appeared on the cover of the New York Post

(b) Tampered image appeared on the Hamevaser front page

FIGURE 1: Examples of fake images of concern of this work.

and videos can seriously harm people involved. Let us think of the audio-video clips of Barack Obama saying derogatory things about several politicians (before revealing the real speaker) [10]. Possibly less dangerous, but still deceptive, is the example in Figure 1(b) where Hamevaser, an Israeli newspaper, digitally removed female world leaders present at Sunday's unity march after Charlie Hebdo attacks.

Given all the negative consequences that the distribution of harmful and/or fake content can cause, the need of dedicated techniques for preserving the dependability of digital media is evident. To this end, the research community recently proposed a novel task within the MediaEval benchmarking initiative (http://www.multimediaeval.org/): Verifying Multimedia Use [11]. This task allowed for the first time to address not only the classical cases where a digital image has been tampered, but also a wider definition of misuse where an image is used in a different place and/or time with respect to the event to which it is associated.

Starting from a preliminary version of the approach of detecting fake content in tweets presented at MediaEval 2016 [12], we develop here the idea in order to face the more complicated scenario of news articles. In this work, images are *real* if they are not tampered and consistent with all facets (e.g., time, location, people involved of the event to which they are associated). *Fake* images, on the other hand, are defined as either tampered (e.g., Figure 1(b)) or miscontextualized (e.g., Figure 1(a)). A miscontextualized image is an image that does not represent accurately the event to which it is associated. Following the definition of the event [13] a miscontextualized image presents inconsistencies with at least one facet of the event, for instance temporal or geographical misplacement or association with wrong event actors (see example in Figure 1(a)).

Our main contribution is twofold: (i) the development of a methodology for discriminating between real and fake images, consisting of image forensics techniques and textual analysis, and (ii) the collection of realistic datasets which are dedicated for testing the applicability of such proposed method in unsupervised scenarios.

It has been acknowledged that there is no single image forensics method that works universally because each method is designed to detect specific traces based on its own assumption; it is therefore wise to fuse multiple output from many forensics techniques [14]. Basing on this acknowledgement, we propose a novel fusion of multiple classical image forensics techniques on the hope to detect various image manipulations. Moreover, we test a recently proposed image forensis tool based on statistical features of rich models [15] and a method based on a Convolutional Neural Network (CNN) [16]. Then, textual analysis is added accordingly to detect real images but being miscontextualized or fake images where traces of manipulation are hidden due to strong postprocessing or poor resolution [17]. The devised methodology is tested on three distinct datasets, the last of which is created by us to investigate the behavior of the algorithm in real-world scenarios and drag insights on the difficulties that might arise.

The rest of this paper is structured as follows. Section 2 outlines the current state of the research for image forensics in online news. Section 3 describes the devised methodology. Section 4, after a general discussion on the three datasets employed, presents the obtained results and a general discussion on the third dataset that is meant to understand performances and weaknesses when the method is applied to real-world scenarios. Finally, some conclusions are drawn in Section 5.

2. State of Art

In literature, there is a prominent research regarding the detection of fake news, especially in social media [18–23]. Similarly, most of the papers that try to discriminate through real and fake multimedia content are focused on content posted on Twitter or other social media. For this purpose the Verifying Multimedia Use task [11] was introduced in 2015, as part of the MediaEval benchmarking initiative, to assess the effectiveness of methods for the automated verification of tweets presenting multimedia content. The definition of the task is the following: "Given a tweet and the accompanying multimedia item (image or video) from an event of potential interest for the international news audience, return a binary decision representing verification of whether the multimedia item reflects the reality of the event in the way purported by the tweet".

Nevertheless, attempts to verify multimedia content in online news have also been made. In [24] a tool that allows verifying the consistency of images used within news articles is presented. To do this, the tool compares images with other visually similar pictures related to the same topic or event, although no analysis of the collected images is performed. In general, image forensics techniques have been frequently employed to solve the overall verification problem. However, it has been shown that these are generally weak for this task. One of the main reasons is that images uploaded online, especially on social networks, are subject to a strong processing, such as compression and resizing. These operations destroy most of the traces left by previous manipulations, thus making extremely difficult their detection [25]. Moreover, image forensic techniques often give inaccurate information in the case of unharmful forgeries, such as the insertion of text or logos, or quality enhancement operations. Therefore, the need of complementing standard forensic features with external information was retrieved online [12, 26] and more specific textual features were proved to be more informative.

Following, we present a brief overview of image forensic and textual analysis techniques that have been used in the literature.

2.1. Image Forensics.

Image forensic techniques traditionally employed by journalists [27] nowadays present several challenges, as getting information such as the date, time, and location an image was taken or getting in touch with the person that published it, are likely impossible or too slow to perform giving fast pace of online news.

Therefore automatic techniques able to assess whether or not a multimedia content is original and to assess which regions are most likely to be modified are needed. Image manipulation is typically classified as either splicing (transferring an object from an image and injecting it into another) or copy-move (copying an object from the same image to a different position). These manipulations normally leave digital traces that forensics methods try to detect. Image retouching, for instance, contrast enhancement, edge sharpening, or color filtering, is not considered in paper since these modifications do not alter semantic content and thus techniques targeting such modifications are not included in our study.

Since JPEG is one of the most common formats of digital images, vast research has focused on several ways to exploit traces left by the JPEG compression process. For instance, different methods have been proposed to determine whether an image was previously JPEG compressed [28, 29] and to discriminate forged regions for double and multiple compressed images [30, 31]. Other techniques to detect tampering exploit inconsistencies in the Color Filter Array (CFA) interpolation patterns [32] and the analysis of quantization tables, thumbnails and information embedded in EXIF metadata to detect nonnative JPEG images [33].

Image manipulations also disrupt Photo Response Non-Uniformity (PRNU), a sort of camera fingerprint that is supposed to be present in every pristine image. PRNU can be therefore used as a useful clue to detect image forgeries [34]. Differently, statistical features of rich models [35] have been successfully exploited by [15, 36] with no a priori knowledge required.

By the advent of deep learning and the amount of available data, image manipulation detection can be solved through Deep Neural Networks (DNNs). The feature extraction task is no more required since DNNs can perform feature extraction and classification through end-to-end process. Highly promising results have been recently achieved, for instance [16, 37–39].

Nevertheless, many of the aforementioned techniques are not always suitable for real cases, where altered images are strongly processed by the social media sites they are uploaded on [25].

2.2. Textual Analysis.

Natural language processing techniques exploit search engines, text corpus visualizations, and a variety of applications in order to filter, sort, retrieve, and generally handle text. Such techniques are typically used to tackle the challenging problem of modeling the semantic similarity between text documents. This task relies fundamentally on similarity measures, for which a variety of different approaches have been developed. Some simpler techniques include word-based, keyword-based, and n-gram measures [40]. An example of a more sophisticated approach is Latent Semantic Analysis [41, 42], where document's topics are learned and expressed as a multinomial distribution over words.

Traditionally, text similarity measurements leverage Term Frequency-Inverse Document Frequency (TF-IDF) to model text documents as term frequency vectors. Then, the similarity between text documents is computed by using cosine similarity or Jaccard's similarity. However, it seems unlikely that many occurrences of a term in a document always carry the same significance of a single occurrence. A common modification is to use instead the logarithm of the term frequency, which assigns a weight to the term frequency [43].

A connected textual technique is the so-called sentiment analysis, which is used to systematically identify, extract, quantify, and study affective states and subjective information. In general, this technique aims to determine the attitude of a writer with respect to some topic or the overall contextual polarity or emotional reaction to a document, interaction, or event. Sentiment polarity text-classification is a challenging task, as determining the right set of keywords is not trivial, although attempts to determine the polarity of sentiments of web pages and news articles achieved a precision varying between 75% and 95%, depending on the nature of the data [44]. The reason behind this considerable gap is that well-written text, such as news articles and descriptions in some official organizational web pages, contain long and complex sentences, which are difficult to deal with. Although the appearance of some keywords is not a direct implication of a sentence to express a positive or negative opinion [45], list of words that could be used for sentiment analysis are available and will be used later on for the textual analysis of the news.
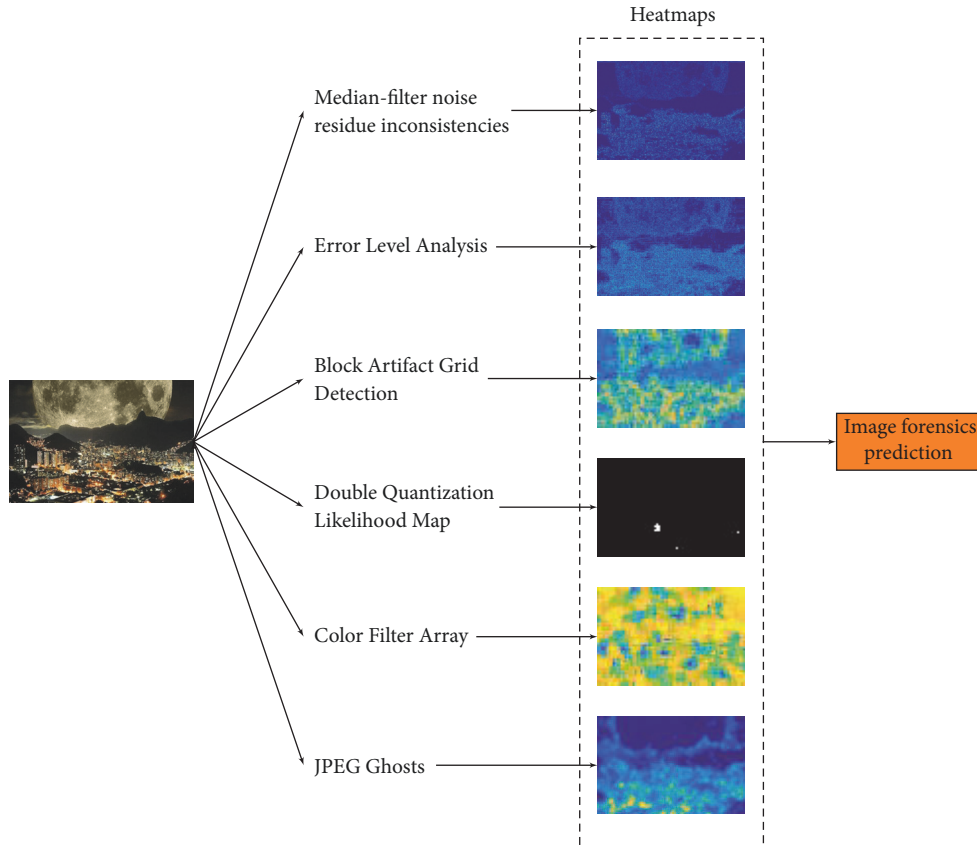
Heatmaps

Median-filter noise
residue inconsistencies

Error Level Analysis

Block Artifact Grid
Detection

Double Quantization
Likelihood Map

Color Filter Array

JPEG Ghosts

Image forensics
prediction

FIGURE 2: Outline of the algorithm for image tampering detection for classical methods.

3. Proposed Approach

The method discussed here was developed to discriminate between real and fake images associated with news articles. The proposed approach uses the framework presented in [12] as a starting point and extends it for the scope of analyzing online news. This extension is motivated by the fact that fake news, typically originated from social networks (which are the concern of [12]), sometimes reaches news outlets and newspapers. This is especially true for high-impact events, where the journalists, under the pressure of being the first to publish the news, might perform only a superficial verification of digital content posted online or even neglect it. Vice versa, unverified news events posted on social networks might come from online news.

Images of concern of this paper are not only those that have been somehow altered, but also images that do not reflect accurately the event described in the news, such as pictures taken at a different time and/or place than the one described, or wrongly depicting other event facets.

Given the duality of the discrimination task, we isolate two subproblems and solve them separately before experimenting different techniques to merge the two methodologies.

3.1. Image Forensics Approach. The first problem consists of deciding whether any manipulation has been performed

on the multimedia content from an image forensics point of view. Three different strategies were applied to tackle this problem, namely, classical image forensics techniques, a method based on statistical features of rich models [15], and recent approach, a method based on deep learning [16].

3.1.1. Classical Image Forensics Methods. The first devised approach applies each of the classical image forgery detection algorithms listed in Figure 2 to generate a heatmap to highlight possible tampering.

Error Level Analysis (ELA) aims at identifying a portion of image that exhibits different compression artifacts compared to the rest of the image. The difference might indicate that the image has been edited. To highlight the modified part, the image is intentionally recompressed at a known quality (95% JPEG quality compression in our implementation) and subtracted from the original image. If the modified area is subject to different compression parameter compared with the remaining area, some irregularities are exposed in the residual image.

Block Artifact Grid Detection exploits knowledge on the JPEG compression. For widely used JPEG compression standard, the blocking processing introduces horizontal and vertical breaks into images, which are known as block artifacts. Although being usually considered a flaw of JPEG, this phenomenon can be used to detect many kinds of manipulations that violate the block structure [46].
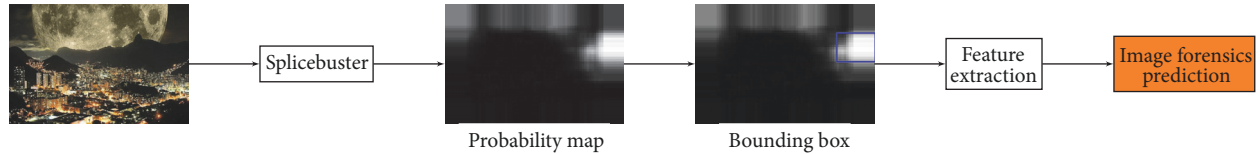
FIGURE 3: Outline of the algorithm for image tampering detection for Splicebuster-based method.

Double Quantization Likelihood Map derived from [30] and can be used to detect tampered regions. The algorithm works by computing the likelihood map indicating the probability for each 8×8 discrete cosine transform block of being doubly compressed for both aligned and nonaligned double JPEG compression. Typically, an image is manipulated (e.g., cropping, copy and paste at a region) and then recompressed. The output image therefore undergoes double JPEG compression, and the probability of being nonaligned is high, 63/64. This makes the method potential for manipulation detection.

Median-filter noise residue inconsistencies detection algorithm [47] is based on the observation that different image features have different high-frequency noise patterns. To isolate noise, a median filtering is applied on the image and then the filtered result is subtracted from the original image. As the median-filtered image contains the low-frequency content of the image, the residue will contain the high-frequency content. The output maps should be interpreted by a rationale similar to Error Level Analysis, i.e., if regions of similar content feature different intensity residue, it is likely that the region originates from a different image source. As noise is generally an unreliable estimator of tampering, this algorithm should best be used to confirm the output of other descriptors, rather than as an independent detector.

JPEG ghosts are based on the premise that, when a splice is taken from a JPEG image and placed in another one of different quality, traces of the original JPEG compression are carried over [48]. In order to detect them, the image is recompressed in all possible quality levels and each result is subtracted from the original. If the image contains a splice, a *Ghost* should appear at the quality level that the splice was originally compressed.

Color Filter Array (CFA) artifacts can be used to localize forged regions [32]. The algorithm is able to discriminate between original and forged regions in an image by making the assumption that the image three colors are acquired through Color Filter Array and demosaicing. By tampering the image, demosaicing artifacts can be destroyed. The detection is done by the usage of a statistical model that allows deriving the tampering probability of each 2×2 image block without requiring to know a priori the position of the forged region.

Each heatmap generated by the described algorithms is then fed to an algorithm that computes the Region of Interest (ROI) of the map, i.e., the region that is more likely to contain tampering, by dividing the image in blocks and finding the one with the maximum variation.

Meaningful statistics (e.g., mean, variance, minima and maxima) are then extracted for that region. These values are finally combined to generate a 40-dimensional feature vector and used either to train or test the manipulation detector.

3.1.2. Advanced Methods. The second approach to identify tampered images is an adaptation of Splicebuster [15]. This detector focuses on splicing detection, which is performed on a single image requiring no a priori information. This method suits well in our scenario as we cannot assume to have any knowledge on the images.

Splicebuster works by extracting local features related to the cooccurrences of quantized high-passed residuals of the image. Then, these features are modeled under two classes, i.e., pristine and tampered, by Expectation-Maximization. The final result is a probability map indicating likelihood of each pixel under pristine model. However, we do not have a ground truth that indicates which area of an image is tampered, and our aim is just to provide a "yes/no" answer with an associated probability that can be then combined with textual analysis. Therefore, we devised the methodology in Figure 3 to convert the probability map to a prediction. Given the probability map extracted from Splicebuster, we try to identify the ROI of the map through bounding boxes. Among these, the biggest one is chosen. Moreover, when the biggest bounding box is smaller than 64×64, we enlarge it to this size keeping the center of the box. The values within the bounding box are then converted to a histogram. Several bin sizes have been investigated; however, results showed how the best results were obtained of 32 bins. Before being fed to a classifier, the histogram is normalized and concatenated to the height and width of the bounding box to get a 34-dimensional feature vector.

The third approach to identify tampered images is based on CNNs, since recently they have been proved to be extremely efficient to solve this type of problem. However, CNNs generally require large labeled dataset for training and thus are hard to be applied in our particular case. Therefore, we leverage the pretrained network in [16]. Although the network allows different image size, we decided to constantly feed 64×64 patches to make it consistent with other tested methods. As can be seen in Figure 4, we adopt the same feature extraction method as we did with Splicebuster in Section 3.1.2: the bounding box of the ROI is computed and then converted to a histogram to be used as a feature vector along with the bounding box size. Also in this case, the final size of the feature vector is 34.

3.2. Textual Analysis Approach. The second problem that we analyze in this paper is how to understand whether an image is coherent with the topic described in the text of the article in which it is inserted.
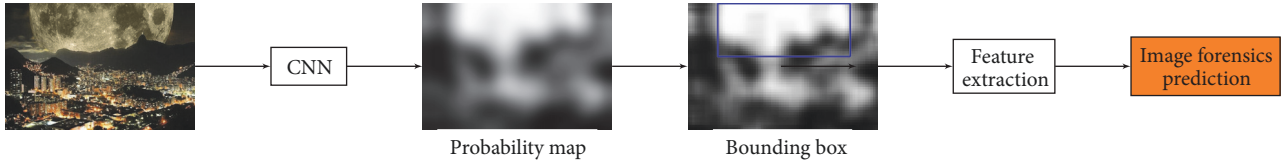
FIGURE 4: Outline of the algorithm for image tampering detection for CNN based method.
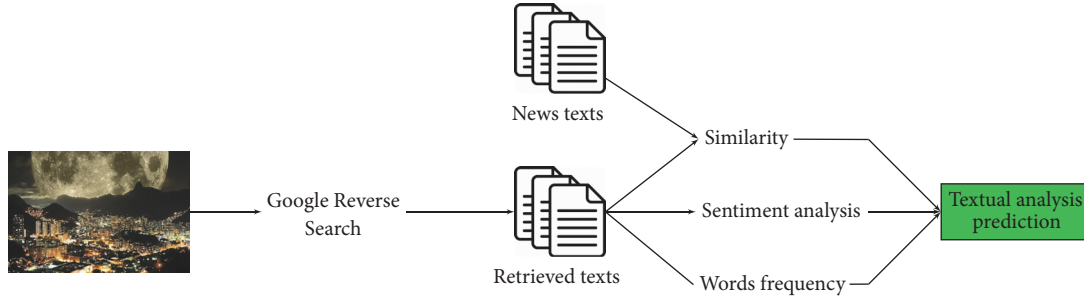

FIGURE 5: Outline of the algorithm for textual analysis.

The approach chosen to tackle this problem allows extracting meaningful values from texts associated with the image under test. This approach, as can be seen in Figure 5, extracts features from two types of documents, namely,

(1) from texts extracted from the news articles related to the event supposedly depicted by the image and

(2) from the texts retrieved online using each image connected to an event as pivot.

The former type of texts are extracted from manually retrieved news articles, which are meant to contain all the words describing the event at stake. By comparing these words with the ones extracted by texts automatically retrieved using the image as pivot, we should be able to detect discrepancies between the event in the news and the story the image is truthfully telling about.

To retrieve text by image, we adopt to use Google Reverse Search. This search engine allows retrieving all online resources that are supposed to contain a given image. Therefore, if the image has been taken before the event to which it is associated, it is likely that articles or resources connected to the first appearance of the image will be collected. Similarly, for tampered images it might be possible to pinpoint pages stating that the image is suspicious.

After retrieving all the text, we proceed to textual feature extraction. First of all the texts associated with the image or retrieved through Google Reverse Search are analyzed to extract the most important words using either TF-IDF, STF-IDF technique or a simple counter.

TF-IDF (Equation (1)), short for term frequency-inverse document frequency, is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus [49]. As the name suggests, this TF-IDF is the combination of two measures: Term Frequency (Equation (2)) used to measure the number of times that term t occurs in document d and Inverse Document Frequency (Equation (3)) used to assess whether the term is common or rare across the set of all documents denoted by D. A highly frequent word is ranked less important if it appears frequently in many documents.

$$\text{TF-IDF}_{t,d,D} = \text{TF}_{t,d} \cdot \text{IDF}_{t,D} \tag{1}$$

$$\text{TF}_{t,d} = f_{t,d} \tag{2}$$

$$\text{IDF}_{t,D} = \log \frac{|D|}{|\{d \in D : t \in d\}|} \tag{3}$$

A common improvement to the previous technique is STF-IDF (sublinear term frequency-inverse document frequency), which assigns a weight given by (4) [43], as many occurrences of a term might not always have the same meaning.

$$\text{STF}(t,d) = \begin{cases} 1 + \log \text{TF}_{t,d}, & \text{if } \text{TF}_{t,d} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

In this task, we also consider to use a simple counter instead of the two commonly used techniques. This was done to evaluate the performance of a rather naïve technique in comparison with more sophisticated ones.

The result of this step, irrespective of which of the three described techniques is used, is a vector of words and word frequencies, as the number of occurrences is normalized by the total number of words. Part of this list of frequencies is used to form the final vector used for classification, to which similarity and possibly sentiment analysis are concatenated.

The similarity is computed by either Cosine or Jaccard's similarity between frequency vectors (A and B) created from the given text and text extracted through Google Reverse Search.

Cosine similarity, which is typically computed for two vectors of frequency of words, is a measure of similarity

TABLE 1: Best performing feature sets for textual analysis.

Feature set	Relevant words extraction	Number relevant words	Similarity measure	Sentiment analysis	Top frequencies
FS1	TF-IDF	All	Cosine		
FS2	TF-IDF	All	Cosine	X	X
FS3	STF-IDF	All	Cosine	X	X
FS4	STF-IDF	Top 100	Cosine	X	X
FS5	Counter	Top 100	Cosine	X	
FS6	TF-IDF	Top 100	Jaccard's	X	X

between two nonzero vectors of an inner product space that measures the cosine of their angle (see Equation (5)).

$$Cos_{A,B} = \frac{A \cdot B}{\|A\| \|B\|} \qquad (5)$$

Jaccard's similarity, computed as in (6), allows indicating the proximity of two sets efficiently.

$$Jacc_{A,B} = \frac{|A \cap B|}{|A \cup B|} \qquad (6)$$

Cosine and Jaccard's similarity can be computed either for the whole vector or only for a subset of the vector, e.g., the top 100 highly rated words.

Finally, basic sentiment analysis techniques are used to analyze people's reaction to the image, which can possibly imply that the image is fake. This is done by analyzing documents retrieved for each image to detect keywords that highlight the feelings toward that image. The computation of sentiment analysis allows extracting, for each text folder associated with each image belonging to an event, three measures: (i) the number of positive words in the text; (ii) the number of the negative words in the text; and (iii) the number of words that are likely to be associated with fake images.

The first two measures are computed by comparing the image's vector of words to a list of words for positive and negative sentiments proposed in [50, 51]. These two measures are useful to determine the general attitude of the writers toward the described events. It was also decided to include a measure that allows to spot keywords that might indicate that an image is fake. Similar to what has been done with the first two measures, the number of words that might indicate that an image is fake is computed. The list of "fake words" used in our work was already used in [12] and contains terms such as "unnatural", "unrealistic", and "retouch".

We conduct a series of testing to determine which combinations of the aforementioned features are worth investigating. During this phase, seven feature sets (FS1 to FS7) are identified. Details on how these feature sets were computed can be seen in Table 1, where *X*s identify the elements being used for a particular set.

In FS3, for instance, STF-IDF scores of all the words are extracted from the retrieved text to form the vector $v1$. The STF-IDF scores of all the words are then computed on the news itself to form the feature $v2$. The final feature vector consists of cosine similarity between $v1$ and $v2$, the

frequencies of top 100 words in the news, and numeric scores from sentiment analysis. As another example, in FS5 the top 100 frequent words and their frequencies are extracted from the retrieved text to form feature vector $v1$. The frequencies of those 100 words are computed on the news itself to form the feature vector $v2$. The final feature vector consists of the cosine similarity between $v1$ and $v2$ and numeric scores from sentiment analysis.

3.3. Image Forensics and Textual Analysis Approaches Combination. Sections 3.1 and 3.2 discussed the approaches devised to solve the problems of detecting tampered and miscontextualized images individually. This section is devoted to the methodology applied to combine these approaches to predict whether the image is real or fake.

We, first of all, train two classifiers separately using respectively image forensics features and textual features. The probability outputs from two classifiers are as follows:

(1) $p(img\_f)$, the probability of an image to be fake from an image forensic point of view (i.e., the image has been tampered) and

(2) $p(img\_t)$, the probability of an image to be fake from the textual analysis perspective, which can include both the scenario of miscontextualized images and tampered images, as already discussed.

These two probabilities are assigned a weight (w_1 and w_2, respectively). The linear combination in (7) represents the probability of an image to be fake: $p(img)$. If this probability is higher than 50%, then the image is classified as fake, otherwise as real, as can be seen in Figure 6.

$$p(img) = p(img\_f) \times w_1 + p(img\_t) \times w_2 \qquad (7)$$

The values of w_1 and w_2 are two percentages that sum up to 100%. In order to evaluate which combinations are more suitable for the task of assessing the authenticity of images within online news, tests are run for all possible combinations of w_1 and w_2. Results obtained with the two tested classifiers, Random Forest and Logistic Regression, will be discussed in Section 4.

4. Experimental Results

In this section we are presenting the results obtained for each of the three datasets described in the followings. The
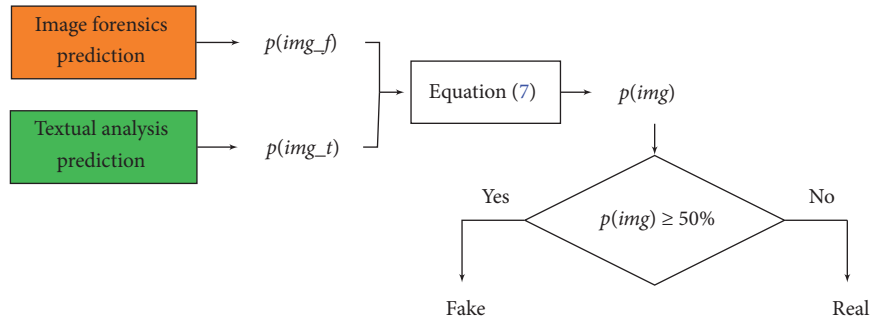
Figure 6: Algorithm for image forensics and textual analysis features combination.

Table 2: Statistics of the collected datasets.

Dataset	Events	Articles	Images	Real images	Fake images
MediaEval2016 (training)	15	170	380	186	194
MediaEval2016 (test)	25	91	98	50	48
BuzzFeedNews	6	56	49	31	18
CrawlerNews	13	130	246	223	23

evaluation is performed in terms of Precision, Recall, and F1-score (see Equation (8)) for two classifiers, namely, Random Forest and Logistic Regression.

$$\text{Precision} = \frac{TP}{TP + FP}$$
$$\text{Recall} = \frac{TP}{TP + FN} \qquad (8)$$
$$\text{F1-score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

4.1. Datasets. The approaches discussed in Section 3 were tested on three different datasets.

In general, as can be seen in detail in Table 2, the datasets are composed of a number of articles and images (either real or fake) grouped by events, such as 2015's Nepal earthquake and 2012's Hurricane Sandy.

MediaEval2016 is the first dataset used for this work. It derives from the 2016 MediaEval competition and in particular from the Verifying Multimedia Use task, where each image is associated with a number of tweets. To make this dataset fit for this task, which is focused on news articles, tweets were discarded and replaced with articles related to the event at stake manually crawled through the Google News' archive.

A second dataset, *BuzzFeedNews*, is created to validate the generalization of the model obtained from *MediaEval2016*. Recently, Buzzfeed became a reliable news outlet highly involved in countering fake news and disinformation online. Therefore, some of the images and news articles that they reported to be fake to sensitize readers to the problem were collected to form this second dataset.

Finally, a third dataset is created to investigate the performance and the weaknesses of the devised approach when applied to an unsupervised real case. This dataset, now on referred to as *CrawlerNews*, has been created as outlined in Figure 7.

To create this dataset, crawls are performed on Google News, a platform that provides useful and timely news in an aggregated fashion, allowing reaching content from many sources simultaneously. The rationale of choosing Google News is due to the fact that many people nowadays turn to such platforms to retrieve information on outbreaking events. Outsell research firm conducted a survey in 2010 revealing that 57% of users turn to digital sources to read news. Among these, more than a half of consumers are more likely to turn to an aggregator rather than to a newspaper site or to other sites [52]. One of the main advantages of Google News is their policy regarding which sources and article to show. In fact, Google News has a list of rules that promotes news sites to provide accountable and transparent information, unique and permanent URLs, and readable content (https://support.google.com/news/publisher/answer/40787?hl=en). These rules should be able to grant a higher level of trustworthiness of the articles.

Since our aim is to analyze news related to high-impact events, we designed a framework to filter news by using five crawlers for five versions of Google News, namely, the Australian, Irish, British, American, and Canadian. These crawlers are responsible for the retrieval of articles (which will be all in English) that are assigned to the Top Stories section of the appropriate version of Google News. To group news related to the same event, we use a similarity threshold on the words in the titles.

Given the news grouped by event extracted by each national crawler, high-impact events were extracted by assuming that such events will be reported world-wide, and thus belonging to the intersection of news acquired by the single crawlers. After a cleaning phase that allows removing duplicated or broken news links, the text of the articles is extracted. Likewise, the news page is parsed to detect and
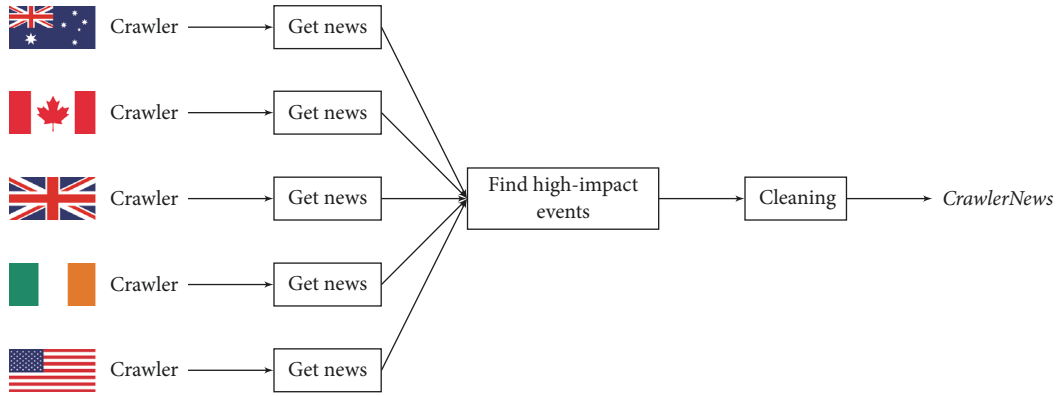
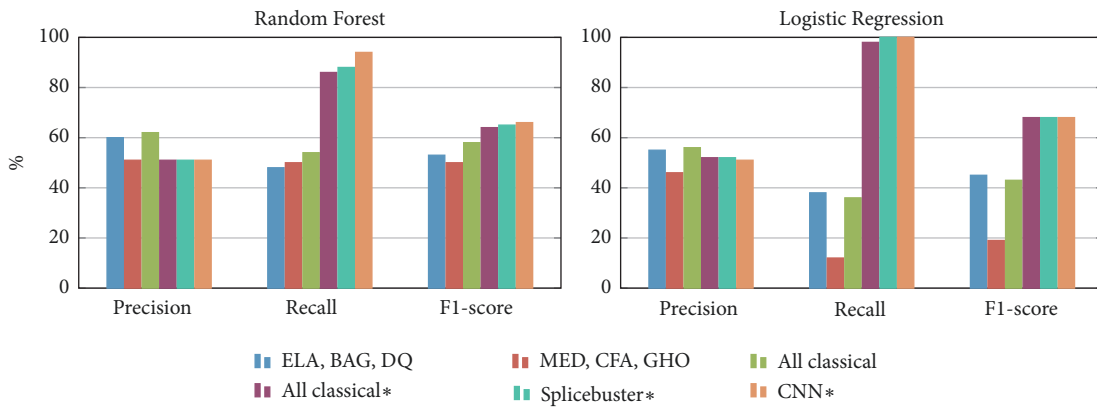FIGURE 7: Outline of the approach used to create the CrawlerNews dataset.

FIGURE 8: Results obtained for forensic features on the *MediaEval2016* dataset. In the legend, the star (∗) means that the new ground truth expressing only whether an image is tampered or not was used for training the classifiers.

extract only images that are relevant to the article, and not advertisements, logos, or images related to other suggested news. This is done by analyzing the position of the image within the text through the HTML tags (only images within the main corpus are kept), the size and some keywords of the image's URL. The thus collected images and texts associated with high-impact news are finally saved to form the *CrawlerNews* dataset. The web crawler was run for about a month, between May and June 2017. The resulting database contains 189 events and around 2500 images. From these only 13 events and 246 images were actually used in this work due to the fact that manually labeling images is hard and extremely time consuming.

In the following sections, results obtained for the three datasets are presented.

4.2. Results on MediaEval2016. This dataset was decisive to be able to evaluate the general performances of the devised approach as well as to run preliminary tests that allowed verifying whether the image forensic features were appropriate and which textual features were the best performing ones.

For image forensic features, the three methods also used in [12], namely, ELA, BAG, and DQ, were compared against the other three methods used for this work and against the usage of all six algorithms for tampering detection (see

Figure 8). In general, using all six methods it was possible to produce a slight improvement. However, an F1-score around 50% cannot be considered really satisfactory. It was therefore hypothesized that this low detection accuracy is due to the fact that the ground truth on which the predictors were trained and tested is irrelevant to our definition about real and fake images. Any original image marked as fake due to its mis-contextualization, and that the image forensic algorithm predicts as real, will in fact downweight the accuracy rate. To validate this hypothesis, a new ground truth, expressing only whether an image is tampered or not, was created. In this case it was possible to produce a remarkable improvement, as can be seen in purple in Figure 8.

This new ground truth was also used to evaluate the performance of the methods based on Splicebuster and CNN. As can be seen in Figure 8, the results of these two methods are only slightly different from the ones of the classical methods, with an improvement in terms of F1-score of 1 and 2% respectively only for the Random Forest Classifier.

The similar trend for the three methods is probably caused by two factors. Firstly, Splicebuster and the CNN were originally designed to provide a tampering map. The prediction based on the tampering map is still an open problem. We resort this problem into local feature extraction on the suspected region. This probably discards global
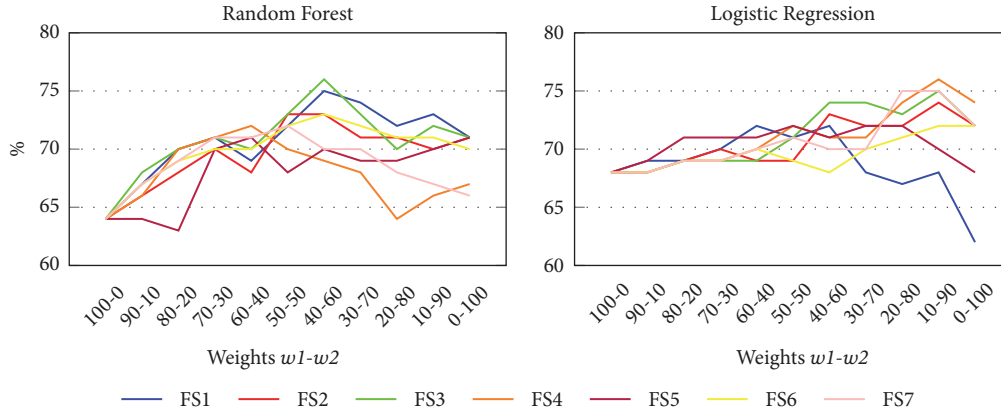
FIGURE 9: Results obtained with classical image forensics features combined with different sets of textual features. Results are reported for different weights for $w1$ and $w2$ as described in Section 3. Vertical axis stands for the accuracy, while the horizontal axis presents different combination strategies.

TABLE 3: Results obtained on *MediaEval2016* for Random Forest Classifier for FS3.

	Classical	Splicebuster	CNN
Image Forensics	64%	65%	66%
Textual Analysis	73%	73%	73%
Combination 40%-60%	74%	74%	73%

TABLE 4: Results obtained on *MediaEval2016* for Logistic Regression Classifier for FS3.

	Classical	Splicebuster	CNN
Image Forensics	68%	68%	68%
Textual Analysis	73%	73%	73%
Combination 10%-90%	75%	75%	74%

information of the tampering map. Secondly, results are very much affected by the quality of online images, which are subject to strong compression and low resolution that might prevent the algorithms from finding tampering traces.

Various tests were run also for textual analysis by combining different textual features. In general, with the analysis of text it was possible to reach an F1-score higher than 70% in most of the cases, which suggests that this type of features might be more suitable to detect fake images.

The results obtained for some of the best performing sets of textual features (listed in Table 1), combined with classical image forensics features using different weights, can be seen in Figure 9.

Random Forest, as can be seen, starts to produce results better than 70% in terms of F1-score from the point in which 70% weight is given to the image forensic features, and 30% to the textual features. The F1-score then keeps rising up until the $w_1 - w_2$ ratio is 40%-60%. At this point a peak of F1-score of 76% is reached for FS3. Also for the other feature sets, results are rather good for this combination ratio. After this point, the F1-score tends instead to slightly decrease.

For logistic regression, the F1-score tends to rise more slowly for most of the textual feature sets. As can be seen in Figure 9, in fact, most of features sets do not reach an F1-score of at least 70% until the textual features are assigned at least

a 40% weight. Nevertheless, the results then tend to remain higher, and the peak, which also in this case is of 76% on F1-score, is reached when image forensics features contribute to the computation of the truth only by a 10%.

Similar observations can be made for the combination of textual features with Splicebuster and CNN based image forensics. In fact, the trend of the curves is analogous to the ones in Figure 9. Tables 3 and 4 summarize the results obtained on *MediaEval2016*.

The obtained results suggest that, despite textual analysis is more suitable for the authenticity discrimination of images, the combination with image forensics features is in general able to outperform their disjoint usage.

4.3. Results on BuzzFeedNews. The main purpose of tests run on *BuzzFeedNews* dataset was to verify that the results just discussed for *MediaEval2016* were not caused by an overfitted training of the two classifiers. In general, it is not possible to find a feature set that worked absolutely better than the others across the datasets. However, FS3 is the one that appeared to have the most consistent behavior both *MediaEval2016* and *BuzzFeedNews* and with the two classifiers used. In Tables 5 and 6 results are therefore presented for this feature set, for the two ratios of $w_1 - w_2$ that appeared to work best for Random Forest and Logistic Regression, respectively: 40%-60% and 10%-90%.

TABLE 5: Results obtained on *BuzzFeedNews* for Random Forest Classifier for FS3.

	Classical	Splicebuster	CNN
Image Forensics	75%	70%	73%
Textual Analysis	64%	64%	64%
Combination 40%-60%	75%	66%	75%

TABLE 6: Results obtained on *BuzzFeedNews* for Logistic Regression Classifier for FS3.

	Classical	Splicebuster	CNN
Image Forensics	77%	76%	78%
Textual Analysis	84%	84%	84%
Combination 10%-90%	86%	87%	86%

Even though feature sets did not behave exactly the same over the two dataset, it is still possible to say that results are good enough to prove that they are not the result of overfitting, as also for this dataset the F1-score is frequently higher than 70% for image forensics, textual analysis, and their combination.

4.4. Results on CrawlerNews. Finally, the methodologies described are applied on an experimental dataset collected through a web crawler as described in Section 4.1. The aim is to try to understand whether the devised method is also suitable to discriminate between real and fake images in an unsupervised scenario. The dataset is analyzed with the best combinations of image forensics and textual analysis. For Random Forest, 85% of the images are predicted as true, while for Logistic Regression the number is slightly lower: around 70%.

To better understand these results, part of the dataset (the 13 events in Table 2) is labeled and analyzed. Tables 7 and 8 show the results obtained for the image forensics, textual analysis, and their combinations on these events. In general, it appears that most of the predictions are actually correct, as most of the images extracted by the crawler are original and in the right context with respect to the news they are associated with. On the other hand, some images predicted as fake by the algorithm are actually real. In order to explain these prediction errors, an analysis of the images and text associated with them has been performed on the extended dataset. It is thus possible to identify three (possibly overlapped) classes of problems that might arise during image authenticity verification in online news, namely, problems related to (i) the extracted events, (ii) the textual analysis, and (iii) the extracted images.

It appears that some events are harder to analyze than others. Among these it is possible to list events related to technology, movies, political events, or gossip. For instance, for the launch of new products or movies or conferences about technological topics the news might contain renderings, graphs, or even logos of the firms at stake that can be misclassified as fake both by the image and textual forensic algorithms. In these cases the misclassification resulting from the image forensic is due to the fact that most images contained in this type of events are computer generated, and not actual pictures. Although some techniques to discriminate between computer generated and natural pictures exists [53–55], they are probably ill-suited for such a vast scenario and integrating them was out of the purpose of this work. For textual analysis the misclassification is caused by Google Reverse Search that, for computer generated images, tends to focus more on recognizing the depicted object and find similar images, than on retrieving the contexts of the image itself.

Political events and gossip might also lead to misclassifications, as they might use satirical images or stock photos of the politicians or people involved. Although this analysis is beyond the scope of this work, it is interesting to note that for political events choosing a particular picture over another can sensibly bias the opinion of a reader. Some of the examples above can be said to belong also to the class of problems related to extracted images, since, as already said, the three classes can overlap. Other types of images frequently misclassified are schema, maps, and images related to the places where the events at stake occurred.

The last class of problems is caused by the quality of the extracted texts and that might lead to misclassification during the textual analysis. In fact, during this phase noise can be introduced as Google Reverse Search interprets an image with the general concepts of *people, officer,* or *location of interest* instead of protagonists and locations of a specific event. For instance pictures (http://www.itv.com/news/meridian/story/2017-06-14/sussex-firm-carried-out-refurbishment-of-grenfell-tower/ [Accessed on Oct. 19 2018]) related to the Grenfell Tower fire returned texts containing the history of the building, and some pictures of Kabul attack (http://www.itv.com/news/2017-05-31/many-killed-and-wounded-in-kabul-car-bombing/ [Accessed on Oct. 19 2018]) were interpreted as *vegetation.*

Other problems might be due to other types of noise introduced by the search. For instance, one of the images (https://www.abc.net.au/news/2017-06-18/portugal-forest-fires-leave-scores-dead/8628896 [Accessed on Oct. 19 2018]) depicting accurately the fire that took place in Portugal on June 2017 was predicted as fake while being original. In this specific case this was due to the fact that many results of Google Reverse Search, although marked as being in English, were actually in Portuguese, which lead to their misclassification.

TABLE 7: Results obtained on *CrawlerNews* for Random Forest Classifier for FS3.

	Classical	Splicebuster	CNN
Image Forensics	92%	88%	92%
Textual Analysis	67%	67%	67%
Combination 10%-90%	85%	80%	85%

TABLE 8: Results obtained on *CrawlerNews* for Logistic Regression Classifier for FS3.

	Classical	Splicebuster	CNN
Image Forensics	93%	93%	93%
Textual Analysis	87%	87%	87%
Combination 40%-60%	88%	87%	88%

On the contrary, an image that was correctly predicted as fake is a demonstrative image, frequently used in association with articles related to the prevention of blood clots during flights. This image has been used in an article (https://www.abc.net.au/news/2017-06-01/what-should-you-do-if-an-incident-happens-on-your-plane/8578102 [Accessed on Oct. 19 2018]) related to an incident on a Malaysia Airlines flight from Melbourne to Kuala Lumpur, where a man threatened to blow up the plane. It is however important to note that, on the article on which the image is posted, the fact that the image is a stock photo is specified in the caption.

In general, although most of the images were correctly predicted as real, the performance of the devised methodology on this dataset is hard to be evaluated since the process of labeling is not trivial. Therefore, the methodology cannot be said to be ready for real-world scenarios, but this experiment is still important to gain insight on possible issues and controversies related to the problem, which might be used in future works to improve the state of the art.

5. Conclusion

The objective of this work is to exploit state-of-the-art techniques to assess image authenticity and relevancy with respect to the news article to which it is associated. This task is extremely important due to the amount of images uploaded everyday online. Those images can get viral within seconds in case of high-impact events. The devised methodology is able to perform rather well on this task, thanks to a combination of image forensics and textual analysis techniques, reaching an F1-score frequently higher than 70%.

Moreover, the analysis performed on a dataset created through a web crawler allows gaining insight on a number of problems that might arise when, instead of using ad hoc datasets, we look into more complex, unsupervised scenarios. Some of these observations suggest the need of more sophisticated techniques to extract text associated with the images, as they are crucial to the correct classification.

In general, the analysis of the last dataset highlighted that methodologies at the state of the art, including the one presented here, present some critical issues when applied to real-world scenarios. To be able to overcome these issues, a possible solution might be the creation of new and bigger dataset with a careful labeling that would also favor a

better exploitation of the power of deep learning based approaches.

Data Availability

Part of the data can be found at the following link: https://github.com/MKLab-ITI/image-verification-corpus/. The remaining will be made publicly available upon request after the acceptance of the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] K. Bulkley, "The rise of citizen journalism," 2012, https://www.theguardian.com/media/2012/jun/11/rise-of-citizen-journalism.

[2] P. Zontone, G. Boato, F. G. B. De Natale et al., "Image diversity analysis: Context, opinion and bias," in *Proceedings of the First International Workshop on Living Web: Making Web Diversity a True Asset*, 2009.

[3] W. Saletan, "The Ministry of Truth. A mass experiment in altering political memories," 2010, http://www.slate.com/articles/health_and_science/the_memory_doctor/2010/05/the_ministry_of_truth.html.

[4] D. Lee, "Deepfakes porn has serious consequences," 2018, http://www.bbc.com/news/technology-42912529.

[5] S. Subramanian, *Inside the Macedonian Fake-News Complex*, vol. 15, Wired Magazine, 2017.

[6] C. Dewey, *Facebook Fake-News Writer: I Think Donald Trump is in the White House Because of Me*, The Washington Post, 2016.

[7] V. Conotter, D.-T. Dang-Nguyen, G. Boato et al., "Assessing the impact of image manipulation on users' perceptions of deception," in *Proceedings of the SPIE 9014: Human Vision and Electronic Imaging XIX*, San Francisco, Calif, USA, 2014.

[8] G. Boato, V. Conotter, F. G. B. De Natale, and C. Fontanari, "Towards multimedia opinion mining," in *Proceedings of the First International Workshop on Living Web: Making Web Diversity a True Asset*, 2009.

[9] N. Golgowski, *Boston Marathon Bombing Bag Men Suing New York Post for Falsely Portraying Them as Suspects*, NY Daily News, 2013.

[10] K. Dold, *Face-Swapping Porn: How a Creepy Internet Trend Could Threaten Democracy*, RollingStone, 2018.

[11] C. Boididou, K. Andreadou, S. Papadopoulos et al., "Verifying multimedia use at MediaEval 2015," in *Proceedings of the MediaEval Benchmarking Initiative for Multimedia Evaluation*, 2015.

[12] Q.-T. Phan, A. Budroni, C. Pasquini et al., "A hybrid approach for multimedia use verification," in *Proceedings of the MediaEval Benchmarking Initiative for Multimedia Evaluation*, 2016.

[13] C. Tzelepis, Z. Ma, V. Mezaris et al., "Event-based media processing and analysis: a survey of the literature," *Image and Vision Computing*, vol. 53, pp. 3–19, 2016.

[14] M. Fontani, T. Bianchi, A. De Rosa, A. Piva, and M. Barni, "A framework for decision fusion in image forensics based on dempster-shafer theory of evidence," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 4, pp. 593–607, 2013.

[15] D. Cozzolino, G. Poggi, and L. Verdoliva, "Splicebuster: a new blind image splicing detector," in *Proceedings of the IEEE International Workshop on Information Forensics and Security, WIFS 2015*, Italy, November 2015.

[16] Y. Liu, Q. Guan, X. Zhao, and Y. Cao, "Image forgery localization based on multi-scale convolutional neural networks," in *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security, IH and MMSec 2018*, pp. 85–90, Austria, June 2018.

[17] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: a survey," *ACM Computing Surveys*, vol. 51, no. 2, 2018.

[18] V. Qazvinian, E. Rosengren, D. R. Radev, and Q. Mei, "Rumor has it: identifying misinformation in microblogs," in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, EMNLP 2011*, pp. 1589–1599, ACL, Edinburgh, UK, July 2011.

[19] J. Ratkiewicz, M. Conover, M. Meiss et al., "Truthy: mapping the spread of astroturf in microblog streams," in *Proceedings of the 20th International Conference Companion on World Wide Web, WWW 2011*, pp. 249–252, India, April 2011.

[20] E. Seo, P. Mohapatra, and T. Abdelzaher, "Identifying rumors and their sources in social networks," *SPIE Defense, Security, and Sensing*, 2012.

[21] X. Liu, A. Nourbakhsh, Q. Li, R. Fang, and S. Shah, "Real-time rumor debunking on twitter," in *Proceedings of the 24th ACM International Conference on Information and Knowledge Management, CIKM 2015*, pp. 1867–1870, Australia, October 2015.

[22] J. Ma, W. Gao, Z. Wei, Y. Lu, and K. Wong, "Detect rumors using time series of social context information on microblogging websites," in *Proceedings of the the 24th ACM International*, pp. 1751–1754, Melbourne, Australia, October 2015.

[23] A. Zubiaga, M. Liakata, R. Procter, K. Bontcheva, and P. Tolmie, "Towards detecting rumours in social media," in *Proceedings of the 29th AAAI Conference on Artificial Intelligence, AAAI 2015*, pp. 35–41, USA, January 2015.

[24] C. Pasquini, C. Brunetta, A. F. Vinci, V. Conotter, and G. Boato, "Towards the verification of image integrity in online news," in *Proceedings of the 2015 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2015*, Italy, July 2015.

[25] M. Zampoglou, S. Papadopoulos, and Y. Kompatsiaris, "Large-scale evaluation of splicing localization algorithms for web images," *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 4801–4834, 2017.

[26] C. Maigrot, V. Claveau, E. Kijak, and R. Sicre, "MediaEval 2016: a multimodal system for the Verifying Multimedia Use task," in *Proceedings of the 2016 Multimedia Benchmark Workshop, MediaEval 2016*, Netherlands, October 2016.

[27] C. Silverman, *Verification Handbook: A Definitive Guide to Verifying Digital Content for Emergency Coverage*, European Journalism Centre, 2014.

[28] W. Luo, J. Huang, and G. Qiu, "JPEG error analysis and its applications to digital image forensics," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 480–491, 2010.

[29] C. Pasquini, F. Perez-Gonzalez, and G. Boato, "A Benford-Fourier JPEG compression detector," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 5322–5326, 2010.

[30] T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 1003–1017, 2012.

[31] C. Pasquini, G. Boato, and F. Perez-Gonzalez, "Multiple JPEG compression detection by means of Benford-Fourier coefficients," in *Proceedings of the 2014 IEEE International Workshop on Information Forensics and Security, WIFS 2014*, pp. 113–118, USA, December 2014.

[32] P. Ferrara, T. Bianchi, A. De Rosa, and A. Piva, "Image forgery localization via fine-grained analysis of CFA artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1566–1577, 2012.

[33] E. Kee, M. K. Johnson, and H. Farid, "Digital image authentication from JPEG headers," *IEEE Transactions on Information Forensics and Security*, vol. 6, 2011.

[34] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Determining image origin and integrity using sensor noise," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 74–90, 2008.

[35] J. Fridrich and J. Kodovský, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 868–882, 2012.

[36] D. Cozzolino, D. Gragnaniello, and L. Verdoliva, "Image forgery detection through residual-based local descriptors and block-matching," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pp. 5297–5301, 2014.

[37] D. Cozzolino and L. Verdoliva, "Camera-based image forgery localization using convolutional neural networks," in *Proceedings of the European Signal Processing Conference (EUSIPCO)*, 2018.

[38] M. Huh, A. Liu, A. Owens, and A. A. Efros, "Fighting fake news: image splice detection via learned self-consistency," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

[39] Y. Wu, W. Abd-Almageed, and P. Natarajan, "BusterNet: detecting image copy-move forgery with source/target localization," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

[40] G. Salton, *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*, Reading: Addison-Wesley, 1989.

[41] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *Journal of the Association for Information Science and Technology*, vol. 41, no. 6, pp. 391–407, 1990.

[42] T. K. Landauer and S. T. Dumais, "A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge," *Psychological Review*, vol. 104, 1997.

[43] H. Schütze, C. D. Manning, and P. Raghavan, *Introduction to Information Retrieval*, vol. 39, Cambridge University Press, 2008.

[44] T. Nasukawa and J. Yi, "Sentiment analysis: capturing favorability using natural language processing," in *Proceedings of the 2nd International Conference on Knowledge Capture, K-CAP 2003*, pp. 70–77, USA, October 2003.

[45] B. Liu, "Sentiment analysis and subjectivity," in *Handbook of Natural Language Processing*, vol. 2, 2010.

[46] W. Li, Y. Yuan, and N. Yu, "Passive detection of doctored JPEG image via block artifact grid extraction," *Signal Processing*, vol. 89, 2009.

[47] B. Mahdian and S. Saic, "Using noise inconsistencies for blind image forensics," *Image and Vision Computing*, vol. 27, no. 10, pp. 1497–1503, 2009.

[48] H. Farid, "Exposing digital forgeries from JPEG ghosts," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 1, pp. 154–160, 2009.

[49] A. Rajaraman and J. D. Ullman, *Mining of Massive Datasets*, Cambridge University Press, 2011.

[50] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '04)*, pp. 168–177, August 2004.

[51] B. Liu, M. Hu, and J. Cheng, "Opinion observer: analyzing and comparing opinions on the web," in *Proceedings of the International Conference on World Wide Web*, 2005.

[52] R. Wauters, "Report: 44% of Google news visitors scan headlines, don't click through," Tech. Rep. 19, TechCrunch, 2010.

[53] Y. Wang and P. Moulin, "On discrimination between photorealistic and photographic images," in *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006*, pp. II161–II164, France, May 2006.

[54] F. Peng, D.-l. Zhou, M. Long, and X.-m. Sun, "Discrimination of natural images and computer generated graphics based on multi-fractal and regression analysis," *International Journal of Electronics and Communications*, vol. 71, 2017.

[55] D.-T. Dang-Nguyen, G. Boato, and F. G. De Natale, "3D-model-based video analysis for computer generated faces identification," *IEEE Transactions on Information Forensics and Security*, vol. 10, 2015.