Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 58 (2019) 217–229

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Detection and veracity analysis of fake news via scrapping and authenticating the web search

Action editor: Alessandra Sciutti

Dinesh Kumar Vishwakarma *, Deepika Varshney, Ashima Yadav

Department of Information Technology, Delhi Technological University, New Delhi, India

Received 18 February 2019; received in revised form 1 June 2019; accepted 21 July 2019
Available online 23 July 2019

Abstract

Social media has become a part of our day-to-day life and has become one of the significant sources of information. Most of the information available on social media is in the form of images. This has given rise to fake news event distribution, which is misinforming the users. Hence, to tackle this problem, we propose a model which is concerned with the veracity analysis of information on various social media platforms available in the form of images. It involves an algorithm which validates the veracity of image text by exploring it on web and then checking the credibility of the top 15 Google search results by subsequently calculating the reality parameter (Rp), which if exceeds a threshold value, an event is classified as real else fake. In order to test the performance of our proposed approach, we compute the recognition accuracy, and the highest accuracy is compared with similar state-of-the-art models to demonstrate the superior performance of our approach.

Keywords: Detection of fake news; Fake news analysis; Fake news in images; Social media; Veracity

1. Introduction

Social media has become an integral part of our life. In the present day's high-tech aura, these platforms are very conducive in disseminating the news to the remotest of the remote areas of the world. Besides, the increasing digitization has substantiated its role in national and international politics. In the modern era of social media, people are frequent in posting images to get more attention, which results in the diffusion of new information over the web. According to Zubiaga, Aker, Bontcheva, Liakata, and Procter (2018) any post which shares multimedia content that does not reliably represents the event that it claims to is considered as a misleading or fake image. On the other hand, the images which are tampered by applying techniques like image splicing, copy-move image tampering, etc. are classified as manipulated images. It creates a big challenge for social media platforms like Facebook, Twitter, etc. to identify the veracity of the massive volume of data posted by the users. Hence, there exists a substantial risk of publishing fake content over social media. Therefore, the veracity of news is the need of the hour, and this research is the move towards addressing this critical problem. Fig. 1 shows some of the examples of fake news spread over Social Media.

Recently, an example was seen during the US presidential election, during which fake news about the Presidential candidates was shared over thousands of times, and the myriad of such fake news spread quickly. Tang, Blenn, Doerr, and Mieghem (2011) showed an example of how content disseminates over social media, whereas (Cheung,

* Corresponding author.
E-mail address: dinesh@dtu.ac.in (D.K. Vishwakarma).

https://doi.org/10.1016/j.cogsys.2019.07.004
1389-0417/

(a) Fake news story spreaded by WTOE 5News during U.S Presidential elections. **(b) Brad Pitt death rumor spreaded by FOX News** **(c) Mark Zuckerberg death rumor**

Fig. 1. Examples of some fake news spread over social media (Courtesy: Facebook).

She, & Jie, 2015) explained how content propagates through the follower network. Tang, Mao, Guessoum, and Zhou (2013) proposed the ISS model, a new rumour diffusion model having three states: ignorant, spreader, and stifler. Along with this, they have also introduced a dynamic friend network model to show how content disseminates through friends' network. This leads to the problem of the user being misinformed. The cumulative effect of misinformed users on social media has a very negative impact (Wessel, Thies, & Benlian, 2016). The spreading of false information also hampers the public emotionally. Nowadays, videos are one of the stepping stone to spread incorrect information due to monetization policy provided by the video platform. YouTube is one of the most popular online video platforms, which allows the user to put their video for monetization. Some video channels are posting information with wrong headlines to influence users from watching their content, thus making them more popular. The video content is entirely different from the headline. Fig. 2(a) shows the hoax related to Sonali Bendre death story posted over YouTube, via "Bollywood Samachar" source as shown in Fig. 2(b), and the response of the public is displayed in Fig. 2(c).

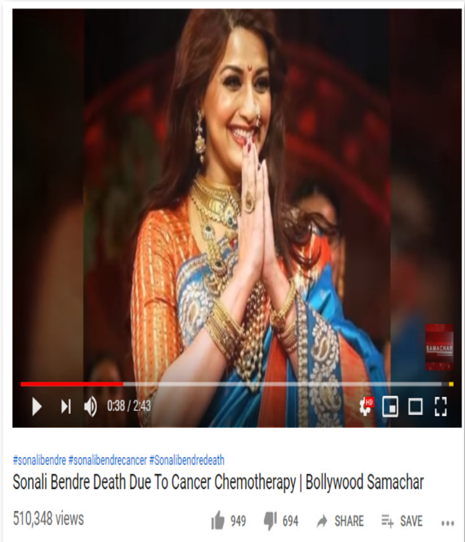
What worsens the situation more is "name theft of authoritative media" (Niu, 2008), which has more effect on the content consumers as it makes them believe the news and also influences their decision making. Many times, trending topics (Aiello et al., 2013) on social media are found to be fake news. This has led to the dire need of a system which separates these noises from the real news events and is also not computationally heavy.

The recent research that has been done in this area is mainly based on two types of algorithms:

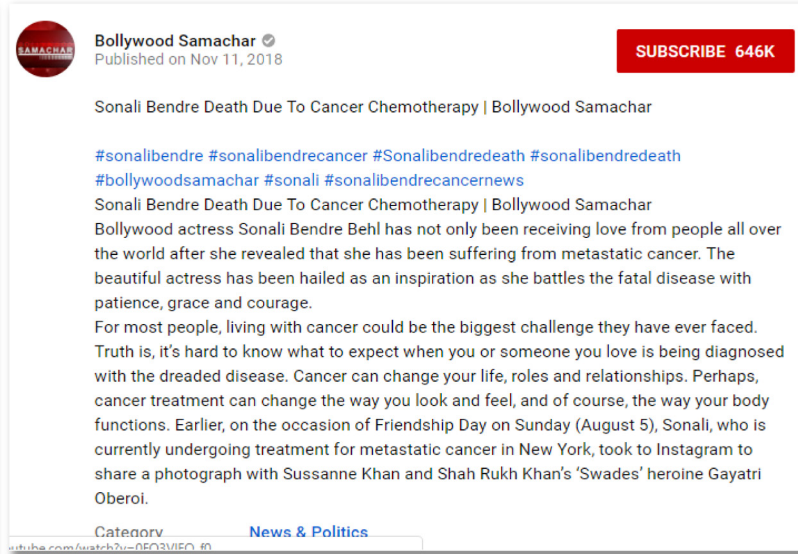
(1) Image-based algorithm, (2) Text-based algorithm. The image-based algorithm extracts various features from the image and trains a model to classify images based on these features. Jin, Cao, Zhang, Zhou, and Tian (2017), proposes five visual features and seven statistical features to validate the authenticity of a news event. Text-based approaches mainly use text pattern and match them with already existing patterns of fake news. They are sometimes referred to as the linguistic approach. A lot of researchers have shifted their interest in credibility detection on posts/tweets using text-based features (Castillo, Mendoza, & Poblete, 2011). Hence, very few studies have been done on images. Our study broadly works on extracting features from the image and text. The proposed system has a very low computational requirement and can be easily implemented and integrated with most of the state-of-the-art systems to stop the propagation of fake news in their vicinity. The algorithm applied in this paper uses various online resources to detect the credibility of the news. For example, news channel like Fox News, CNN, NDTV, AAJTAK, and newspapers websites like The Hindu, The Times of India, The Washington Post, New York Daily News are used to detect the credibility of the news. The results of the algorithm are highly dependent upon the sources mentioned above, whose trustworthiness is mandated by their good name in the market.

Our significant contributions to this area and the key features of the proposed technique can be summarized as follows:

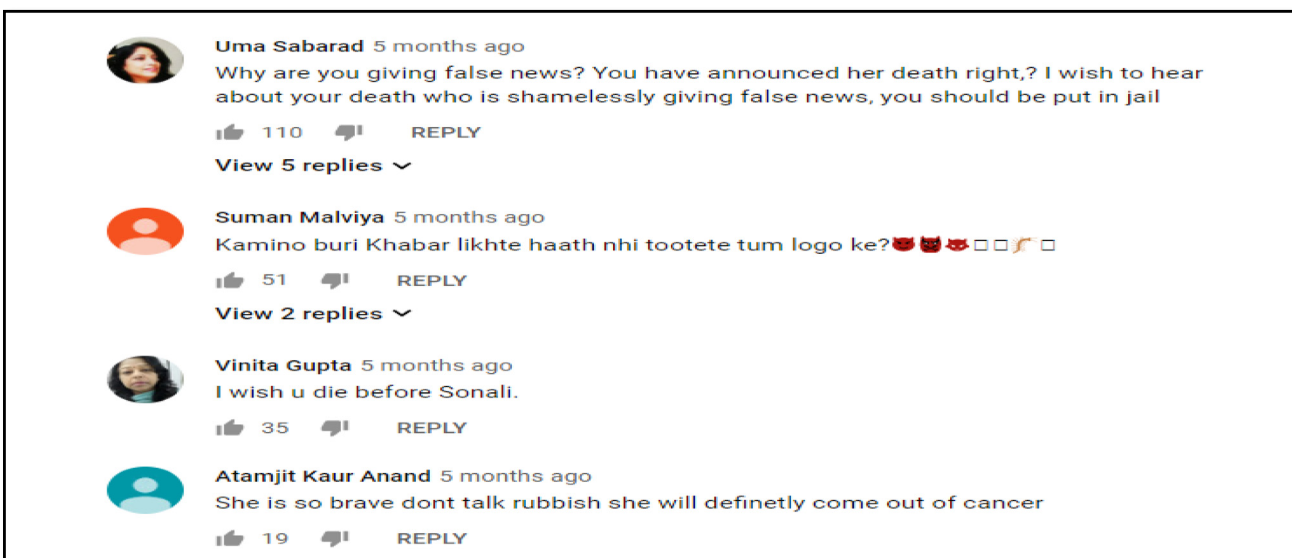
- We propose a novel fake news authentication system for detection of fake news on social media platforms like Facebook and Twitter. This model comprises of four integrated units, namely, text extraction from image, entity extractor, scraping the web and processing unit.
- We show that our system can reliably detect fake news on various social platforms.
- We removed false positives by calculating the similarities between the title of the page and entities extracted from the image and further looking for selected keywords.



(a) Death Rumour of Sonali



(b) Source of Rumor



(c) Public Comments on rumor

Fig. 2. One of the hoaxes related to Sonali Bendre's death spread over social media.

- To validate the proposed fake news detection framework, the dataset is collected from popular social networks like Facebook, Twitter, etc.
- The system, after a feasibility study, has been designed and developed from the ground. The system exploits available technologies for text analysis and web scraping.
- We also perform a comparison based study with other state-of-the-art for instantaneous rumour detection.

The list of frameworks and libraries used are: (a) Natural Language Toolkit (NLTK) (Steven, Loper, & Klein, 2009), a Python package for natural language processing, (b) Selenium (Holscher, 2015), a suite of tools to automate web browsers, used for scraping Google search results, (c) Scrappy (Hoffman, 2013), a web scraping framework of Python.

The rest of the paper is structured as Section 2 gives the recent state-of-the-art in the fake news detection; Section 3 presents the proposed architecture and its functioning; Section 4 explains the process of experiments to evaluate the performance of the system and presents the final results. Finally, Section 5 gives the conclusion and future work.

2. Related work

Credibility evaluation of the information is one of the most popular areas before fake news/rumours detection, which has been already studied (AlRubaian, Al-Qurishi, Al-Rakhami, Rahman, & Alamri, 2015; Byungkyu, O'Donovan, & Höllerer, 2012; Castillo et al., 2011; Li, Dai, Ming, & Qiu, 2016). In the recent era, false information is spreading over social media very fast, and the detection of false news is becoming more challenging. Hence, automatic fake news detection is one of the emerging research fields, which also relies on assessing the credibility of a message. Very few studies have been done in automatic fake news detection (Zubiaga et al., 2018). The recent research that has been done in the area of fake news detection have been mainly on images and text. The image-based algorithm extracts various features from the image and trains the model to classify images based on these features. Text-based approaches primarily use text pattern and match them with already existing patterns of fake news. These are sometimes referred to as the linguistic approach.

(a) Fake News Detection on Images

Very few studies have been done on fake news detection on Images. The most recent work on images in the field of fake news detection is given as follows:

Jin et al. (2017) state that the images have a distinctive distribution pattern for real and fake news. They propose a set of visual features for news verification, which is extracted from visual image content and reveals hidden characteristics of image distributions in news event. Apart from the visual features, it also proposes several statistical image features used to summarize image statistics. Visual features, in combination with statistical features, are used for news authentication. Elkasrawi, Dengel, Abdelsamad, and Bukhari (2016) proposed an algorithm which uses Google reverse image search to find the similar images and applies K-means clustering (based on publishing day) to get the general overview of how the images were used across time. To detect the alteration in the image (based on the assumption that the original image exists somewhere on the web), it uses techniques like edge detection, scaling, colour conversion. Based on alteration in the image, the difference in publishing date of a news event, and that of the first occurrence of the image on the web (image occurred before news event), it labels the events as fake or real. Jin, Cao, Jiang, and Zhang (2014), builds a content-based three-layer hierarchical credibility network to evaluate the credibility of news. The three layers are the message layer, sub-event layer, and an event layer. Pasquini, Brunetta, Vinci, Conotter, and Boato (2015) verifies the consistency of visual and semantically similar images used within different news articles on the same topic, whereas (Rashed, Renzel, & Klamma, 2011) makes a Multimedia Quality Profile consisting of a combination of metadata drawn from fake multimedia management and detection system, visual feature extraction modules, and a trust management module to detect fake multimedia.

In addition to this, the images found over the social web are also likely to be the result of image tampering. Many researchers have shown their keen interest in multimedia forensics methods for detecting tampered images on the web. Past studies provide a survey of several multimedia forensics algorithms for identifying tampered images. Zampoglou, Papadopoulos, and Kompatsiaris (2017) provides a detailed review analysis which includes the robustness, weaknesses, and the future aspects of splicing localization algorithms for image tampering. Nowadays, it has been observed that the copy-move image tampering detection is one of the crucial aspect, as it can be one of the stepping stone to encourage misleading content over the web. Hence, Ferreira et al. (2016) combined different characteristics of copy-move detection approaches by modelling the problem on a multiscale knowledge behaviour space. Dadkhah, Manaf, Hori, Hassanien, and Sadeghi (2014) proposed an effective self-recovery algorithm based on singular value decomposition for tamper detection. Experimental results show that the proposed approach is efficient for tamper detection as it can achieve a tamper detection rate of more than 99%.

(b) Fake News Detection on Text

Many authors have worked on veracity classification problem (Ahmet et al., 2017; Chang, Zhang, Szabo, & Sheng, 2016; Zhang, Zhang, & Li, 2015). The most recent work on the text in the field of fake news detection are given as follows:

Alrubaian, Al-Qurishi, Hassan, and Alamri (2018) assess the problem related to information credibility on Twitter. They have proposed an automated classification system, including four major components: (1) The reputation-based technique, (2) A credibility classifier engine, (3) A user experience component, and (4) A feature rank algorithm. Novelty and pseudo feedback (PF) based features have been introduced by Qin, Wurzer, Lavrenko, and Tang (2016) to detect rumours on early basis, along with features based on the presence of several URLs, hash-tags and user-names, POS tags, punctuation characters as well as eight different categories of sentiment and punctuation emotions. Many authors have worked on veracity classification task (Chang et al., 2016). Vosoughi (2015) introduced three sets of features related to linguistic, user-oriented, and temporal propagation. The Twitter dataset has been used for evaluation. Study reveals that the best performing features were those in the temporal category. Sarcasm is also one of the crucial issues over social media. Bouazizi and Ohtsuki (2016) assessed the problem related to sarcasm on twitter using pattern-based approach and introduced four sets of features that cover the different type of sarcasm and classified tweets as sarcastic and non-sarcastic.

Social media is an open community where anyone can create their content, without any check on its veracity. Also, data present on social media is highly heterogeneous (Pang et al., 2015). Though, many credible sources are there whose integrity cannot be questioned (Niu, 2008) and the content produced by them is verified and double checked. Inspired by these ideas, we exploit this property in our work. We are proposing a novel approach to tackle the problem of fake news. To the best of our knowledge, this is the first attempt to solve the problem with such readily available technologies which is not computationally expensive.

3. Proposed architecture

The flow diagram of the proposed algorithms is shown in Fig. 3. It can be seen how all smaller units are connected, and finally, an overall system is developed. The purpose of the proposed system is to analyze the veracity of the news events that are floating in the form of images in social media. The framework is composed of four basic units: (i) Text extraction from image (ii) Entity extractor (iii) Processing the Web (iv) Processing Unit. In the following section, the details about each module are discussed in depth.

(a) Text Extraction from Image

The input image goes through a series of transformation which facilitates the process of text extraction. The first module, "Text Extraction from Image," performs the function of extracting text from the image. Here we use the method proposed in Chidiac, Damien, and Yaacoub (2016) for detection of text region and then with the help of optical character recognition (OCR), the text is extracted from images. The key steps of the algorithm are: Firstly, Maximally Stable Extremal Region (MSER) detection is used to locate every text location comprised of text with various fonts and size. Secondly, Maximally Stable Extremal Region (MSER) enhancement is used to make boundaries of letters more identifiable. Thirdly, Stroke Width Detector is applied for the detection of stroke width of characters. For this purpose, ray vectors are calculated using Eq. (1). Finally, filtering is done to

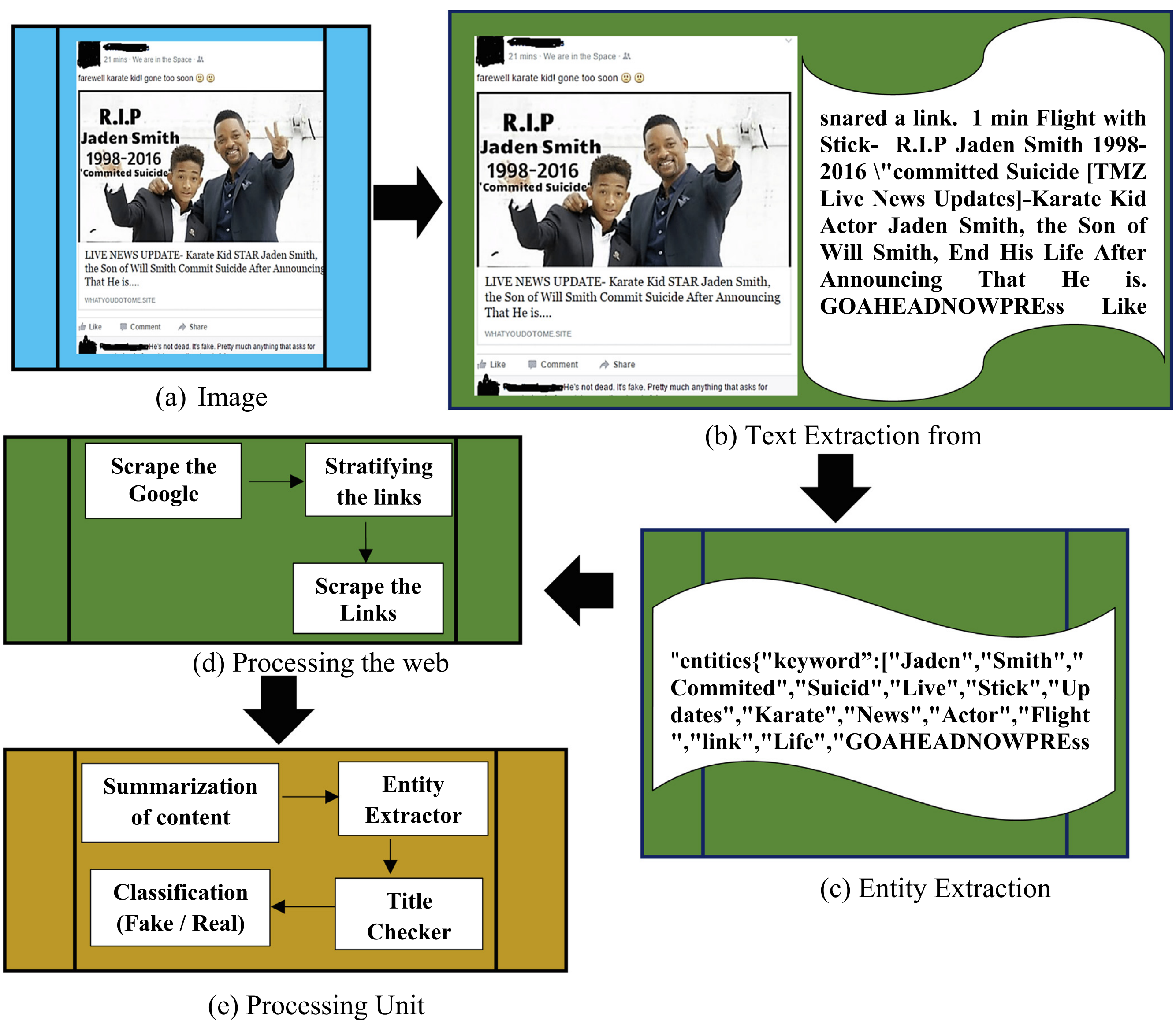


Fig. 3. Underlying Architecture of the proposed system.

eliminate the area which is unlikely to contain text characters. Text region components recognized in the above process goes through Optical Character Recognition module to recognize text from the image. In this paper, we considered only the English language.

$$Rx = Px + n\cos(\theta), \quad Ry = Py + n\sin(\theta) \tag{1}$$

where ray vector $r = [R_x, R_y], p = [P_x, P_y]$ represents the boundary pixel position, n is the iteration index, and θ represents the gradient direction. Eq. (1) is used to find the boundary pixel in the edge image by increasing the value of index n. If the direction of two boundary pixels is the opposite, then the ray is included, else discarded.

(b) Entity Extractor

The Second module, "Entity Extractor," is responsible for the extraction of entities from text. In the following section, the process of entity extractor is explained in brief. Initially, the extracted text from the image is processed to fetch the various entities from it. Then each entity goes through the process of text cleaning, which is further responsible for the following: (1) striping of all non-alphabetic characters, (2) removing multiple occurrences of the words, (3) checking whether the word is a valid English word, (4) checking each word for spelling errors with 1-edit distance, (5) removal of any media house name or newspaper name to remove bias. Fig. 4 shows a working example of entity extraction.

(c) Processing the Web

In the third module, the string of extracted entities are searched on Google, and links to the search result are collected, which are further scraped for their content. Hence, in this module, the Google search results are fetched and are further categorized into a reliable or unreliable link, followed by scraping the content of the reliable links. To this aim, web scraping techniques are applied. Algorithm 1 describes the process, and it returns the dictionary with links as the key, and content as the value. The following steps performed in this module are described in detail:

- Scraping Google is the first step in this module, which searches a specific string on Google search engine and scrapes the results. To this aim, we use Selenium, a portable software testing framework for web applications.
- Stratifying the links is responsible for classifying the links into reliable or unreliable based on a list compiled by us.
- Finally, scraping the links are concerned with scraping the content of each reliable link. To this aim, we use the already-available tool Scrappy, a framework written in Python for the same.

Algorithm 1.

```
procedure ProcessTheLink (text)
   L ← GetGoogleSearchResults (text)
   R ← GetReliableLinks (L)
   For each link in R
      Content ← ScrapeLink (link)
    return Content
end procedure
```

(d) Processing Unit

The fourth module, "Processing unit," is the final module in the proposed system and is concerned with the classification of the event. The classification is done into two

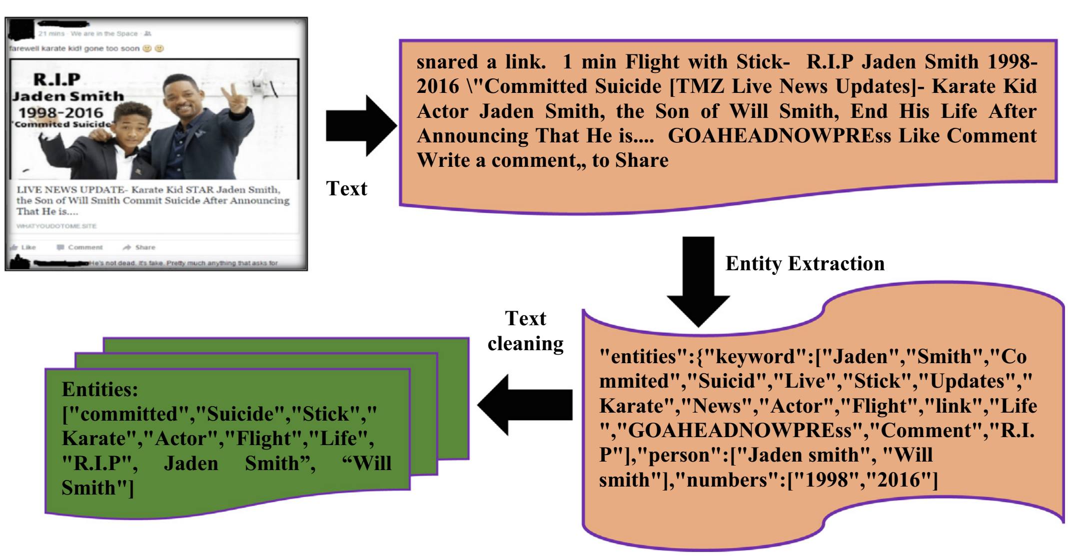


Fig. 4. Entity extractor Process.

categories "fake" or "real". In the following text, submodules are described with their working:

- Summarization of content is concerned with producing a concise summary of the content of the web pages. To help us in this process, we use Python's natural processing tool (NLTK). For obtaining a summary, the extraction parameter has been tuned to suit our need. Based on the values obtained at each iteration, the best results were achieved by limiting the content of the summary to 40–65 words. Moreover, only reliable links go through this process.
- Entity extractor is concerned with the extraction of entities from the summary produced by summarization of content.

Title Checker is responsible for calculating the percentage match of entities extracted from the image with that of the title of each reliable link. Furthermore, if a match is above a threshold value and a particular keyword is found in the title, then the link is not considered as reliable. Algorithm 2 describes title checker process where three parameters (title, query, bag of words) are passed to the title checker procedure.

Algorithm 2 (*Title Checker*).

procedure TitleChecker (*title*, *query*, *bag of words*)
    P ← **percentage match** (*title*, *query*)
    If P > 0.30 and the title has "bag of words"
        return False
    Else
        return true
    End If
End procedure

Classification is the submodule where the final decision is taken based on the value of reality parameter. The content of the reliable links goes into a summarizer, which summarizes the content. The summary goes through entity extractor, and then the percentage match is found between extracted entities from summary with query searched on Google. If the match comes out to be zero then the link shift to unreliable link, else it stays as reliable. Furthermore, the content of the reliable link further goes through title checker, which if return false, the link is considered as unreliable. The number of reliable and unreliable links are then used to find the reality parameter, which is used to classify news as fake or real. Reality parameter value has been calculated using Eq. (3) to classify the news as fake or real. If Rp ≥40 then the news is classified as a real, else it is classified as fake. Algorithm 3 describes the final module:

Algorithm 3 (*Fake News Classification*).

procedure Classifying Real And FakeNews (reliableLinks, Content, query, totalLinks)
    totalReliableLinks = 0
For each link in reliableLinks
If (TitleChecker (Content [link]) and SummaryMatch (Content [link], query) > 0.0)
        totalReliableLinks = totalReliableLinks + 1
    End If
End For
    Rp = (totalReliableLinks/totalLinks) * (1 0 0)
    If Rp ≥ 40
Classify as real
Else
    Classify as fake
End If
End procedure

4. Experimental state-of-the-art

The credibility of the multimedia content on social media is a new and emerging problem, and there are very limited datasets available for classification of an image as fake or real. One of the publicly available dataset (PHEME) of rumors and non-rumors is provided by Zubiaga, Liakata, and Procter (2016). The dataset includes a collection of 1972 rumors and 3830 non-rumors associated with five breaking news stories from Twitter. Boididou et al. (2018) collected dataset from VC-MediaEval 2015 task, which consists of tweets related to 17 hoaxes including 193 real images, 218 fake images, and two cases of misused videos. Another dataset related to rumor and non-rumor is provided by Vosoughi, Mohsenvand, and Roy (2017). This dataset contains 209 rumors, including 938,806 tweets concerning real-time events (2013 Boston Marathon bombings, 2014 Ferguson unrest, and the 2014 Ebola epidemic, etc.). Table 1 describes some of the publicly available datasets for fake news detection. Moreover, Zhang and Ghorbani (2019) presented a survey which discussed the publicly available datasets for fake-news analysis.

Hence, to test the performance of the developed algorithm, a dataset of thousands of images are collected from Google Images, the Onion, and Kaggle maintaining a balance between fake as well as real news images. The sample images of datasets are as shown in Fig. 5. The news is further divided into three categories: (1) Local/regional news, (2) National news, (3) International news. The cases per category are shown in Table 2. For the list of verified content producers, we have manually compiled the maximum possible name into a common file.

Table 1
Dataset Details for fake news analysis.

References	Dataset description	Input data type
Zubiaga et al. (2016)	PHEME dataset includes a collection of 1972 rumors and 3830 non-rumors associated with five breaking news stories	Text
Boididou et al. (2018)	Data is collected from the list of 17 news stories events in VC MediaEval 2015 containing 193 real images, 218 fake images, and two cases of misused videos	Text and Images
Vosoughi et al. (2017)	Dataset is comprised of 209 rumors, including 938,806 tweets, concerning real-time events	Text
Shu, Wang, and Liu (2017)	Fakenews Net dataset contains 211 fake news and 211 true news that are gathered from BuzzFeed.com and PolitiFact.com	Text
Horne and Adali (2017)	The BuzzFeed election dataset contains 36 real news stories and 35 fake news stories during the 9-months before the US Presidential Election	Text

(a) Number of Links

To calculate the optimum number of links required to classify the image, we conducted a series of iterations. Table 3 shows the result of those iterations on all three types of news, i.e., regional, national, and international. From Table 3, we choose an optimal value for the number of links to be 15.

(b) Reality Parameter

To calculate the effective value of Reality parameter, we conducted a set of iterations on our dataset for different values of reality parameter. Table 4 shows how the system performs on different values of Rp for national and international news. In the case of local news, the system is not able to classify local news events as real or fake, as most of the prominent content producers do not cover this news. To calculate the accuracy (Acc) of the system, we have used Eq. (2).

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} * (100) \tag{2}$$

where TN, TP, FN, and FP are the True Negative, True Positive, False Negative, and False Positive, respectively. In our analysis, we also found that for some cases, the value of reality parameter comes out to be exactly 40% as the Google search index of unreliable sites are better than the reliable sites, and they have better rank on Google search result. This resulted in the inclusion of those unreliable sites in our set of links, thus shifting out the reliable links from our set. The detailed analysis of factor affecting Google search result is listed in Zhu and Wu (2011). To calculate the reality parameter, we have used Eq. (3). The proposed system gives the best accuracy at 40%.

$$Rp = \frac{Number\ \ of\ \ reliable\ \ links}{Total\ \ number\ \ of\ \ links} * (100) \tag{3}$$

During our experimentation, we have found that sometimes, credible media sources cover the fake news stating it as a hoax to aware the readers. Table 5 shows how different fake news is covered by credible news media, but their context is different, and their purpose is to make their readers aware of the fake news events. We have analyzed the news titles and found that there exists a pattern of specific keywords in the title encountering fake news. One of the fact-checking website Snopes.com[1] has been analyzed to identify the pattern of news titles covering the fake news. We have manually extracted the list of 10 frequently used keywords and have included them in our bag of words. The keywords are: Hoax, Hoax fools, False News, Fake, Fake news, Fake death, Rumors, False, Death Hoaxes, Falsely. These bag of words have incorporated in our algorithm to remove any false positives we might get due to such cases.

[1] https://www.snopes.com/fact-check/celebrity-death-hoaxes/.

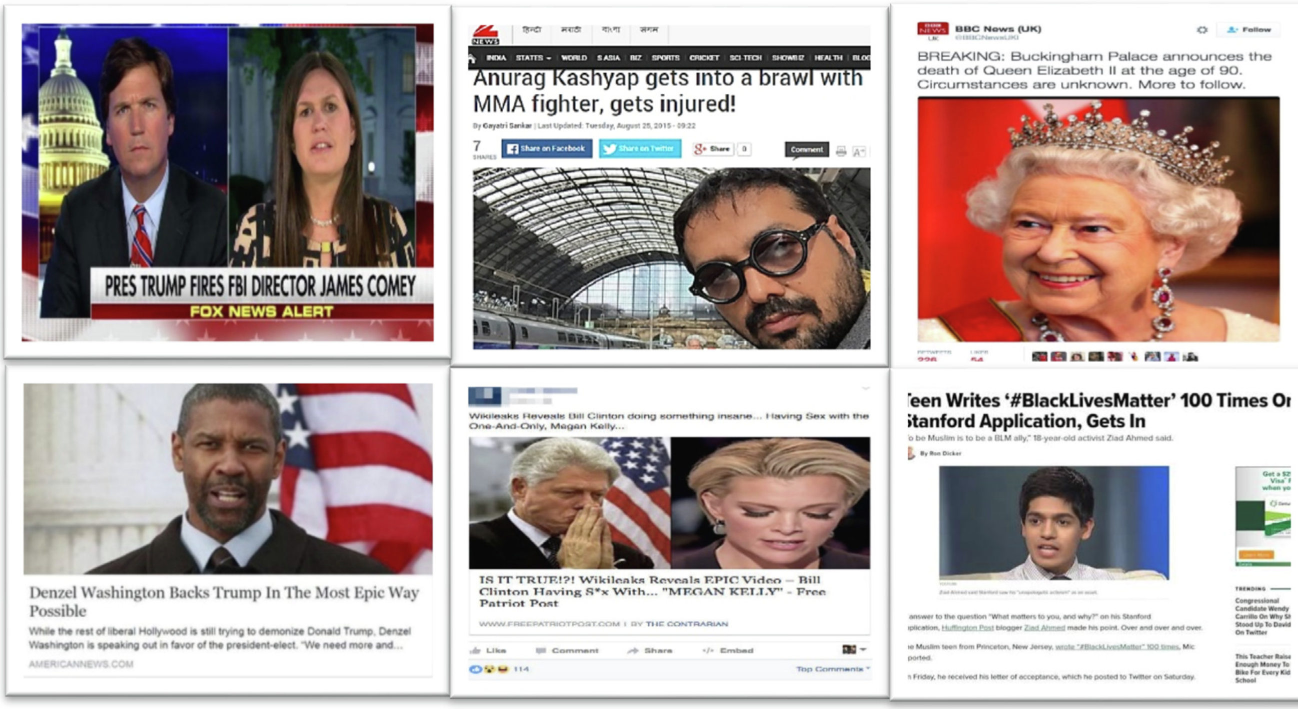


Fig. 5. Sample images of the Dataset.

Table 6 shows the percentage match of page title with news type (national and international news). The minimum value of the title match is chosen as the threshold which is set to 0.30. News events include both fake and real events. Cases, where 0% match is found, are not shown in the table for both fake (cases where news is covered by credible sources to alert their readers) and real news.

Sometimes due to miscommunication, or under stress or excitement, credible media or newspapers also publish fake news. But after a while, realizing their mistake, they take down the respective web page. But by the time they do that, Google has already crawled and indexed their page, and it starts showing in the Google search results. Now when we search any of such event, it gives us a credible source as a result, but there is no relative content present in that link. Due to this reason, we kept a check on summary match to be more than zero. Table 7 shows the summary match for real news.

(c) Effectiveness of fake news detection

To evaluate the effectiveness of our proposed method, a comparative analysis is outlined with three state-of-the-art

Table 3
Variation in the number of reliable links.

NLC	NRLFNE	NRLRNE
0-5	1-3	2-5
6-10	-	2-10
11-14	1-5	2-11
15	1-5	2-12
16	1-6	2-12
17	1-6	2-12
18	1-6	2-13
19	1-6	2-13
20	1-6	2-13

NLC: Number of Link under consideration, NRLFNE: Number of reliable links for fake news events (min-max), NRLRNE: Number of reliable links for real news events (min-max).

Table 4
Effect of Reality Parameter on Accuracy.

Rp	≥80	≥66	≥53	=46	≥40
ANIN	68	73	77	78	85

Rp: Reality parameter (Percentage), ANIN: Accuracy on National and International News (Percentage).

Table 2
Category-wise distribution of images.

S. NO	News category	Cases per category (Total_cases = 1000)	#Fake images	#Real images
1	Local/Regional News	150	70	80
2	National News	350	170	180
3	International News	500	250	250

Table 5
Fake News covered by credible media to aware users.

List of news	Title found	Keyword
Brad Pitt died	Brad Pitt is NOT dead as vile online hoax fools fans with virus amid divorce from Angelina Jolie	Hoax fools
Women killed, Black lives matters	The story of 19 white women killed by Black Lives Matter supporter is fake news	Fake news
Women defecating boss	How fake news story of woman defecating on boss' desk after hitting $3m jackpot fooled thousands after going viral	Fake news
Facebook to start charging	Will WhatsApp and Facebook start charging? Latest scam tells users to pass on chain messages to avoid costs	Latest scam, scam
Jaden Smith died	Jaden Smith STILL not dead after vile 'suicide' hoax continues to baffle fans online	Hoax

Table 6
Title match with types of news.

Type of news	National news (Min-Max)	International news (Min-Max)
Title Match (%age)	0.31–0.71	0.30–0.64

rumor detection methods: Qin et al. (2016), Liu, Nourbakhsh, Li, Fang, and Shah (2015), Yang, Liu, Yu, and Yang (2012). The comparison has been done for instant rumor detection with respect to the following five factors: (1) Proposed Features, (2) Input Data Form, (3) Data Transformation, (4) Dataset, and (5) Accuracy. From the analysis, it has been observed that our proposed approach outperforms the above state-of-the-art in terms of computing accuracy. Table 8 shows the comparative analysis of our technique with the other state of the art approach and the accuracy achieved through the proposed method is highlighted with the bold letter.

From the above analysis, it has been observed that the best accuracy is achieved via our proposed approach when the Rp (Reality parameter) value is 40% which outperforms the other state-of-the-art methods for instantaneous fake news detection. Qin et al. (2016) proposed novelty and pseudo-feedback (PF) features, and along with it, used other features introduced by the two baselines (Liu et al., 2015; Yang et al., 2012) to detect rumor instantly. They have collected data for rumor and non-rumors from Sina Weibo and Xinhua, one of the most popular news agency in China. The detailed description of the same is shown in Table 8.

Additionally, the performance of proposed algorithm is measured on the datasets used by different state-of-the-art (Horne & Adali, 2017; Shu, Wang, & Liu, 2017; Zubiaga et al., 1610) as shown in Table 9. The evaluation metrics used for measurements are Accuracy (Acc), Precision (P), Recall (R), and F1-scores. All the experiments were performed using five-fold cross-validation settings, where the final accuracy is computed by averaging the values across each of the five folds.

From the Table 9, the evaluation metrics parameter shown in bold letter are achieved by the proposed algorithms and the value of these parameters are higher than the others state-of-the-arts. Hence, it can concluded that the proposed algorithms outperforms in comparison with other state-of-the-arts on similar datasets.

Table 7
Showing summary match results of news types.

Type of news	Local news	National news (Min-Max)	International news (Min-Max)
Summary Match (%age)	–	0.274–0.485	0.23 – 0.41

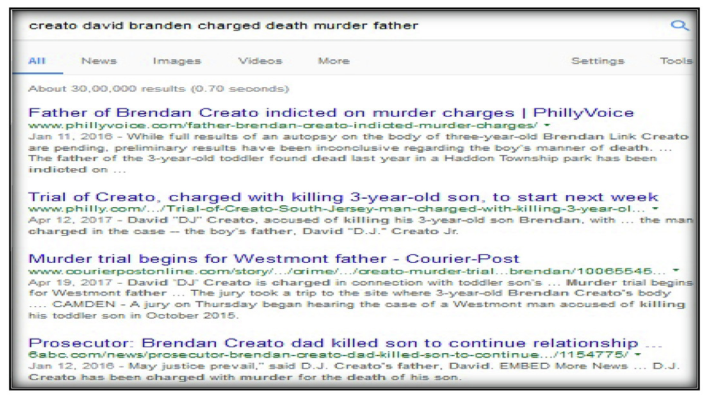
Table 8
Comparative analysis with the other state-of-the-art.

Methods	Proposed features	Input data form	Data transformation	Dataset	Acc
Qin et al. (2016)	Novelty and pseudo feedback (PF) including Punctuation, POS, Sentiment, Emotion, Social Media Length, URLs Novelty.	Text	NA	Xinhua News Agency 2, Sina weibo.	75%
Liu et al. (2015)	Number of URLs, hash-tags and user-names, POS tags, punctuation characters.	Text	NA	Twitter	62.27%
Yang et al. (2012)	Number of URLs, hash-tags and user-names, POS tags, punctuation characters.	Text	NA	Sina Weibo	60.21%
Our method	Reality parameter (Rp)	Images	Image to Text	Google Images/ Kaggle/The Onion	**85% (when Rp = 40)**

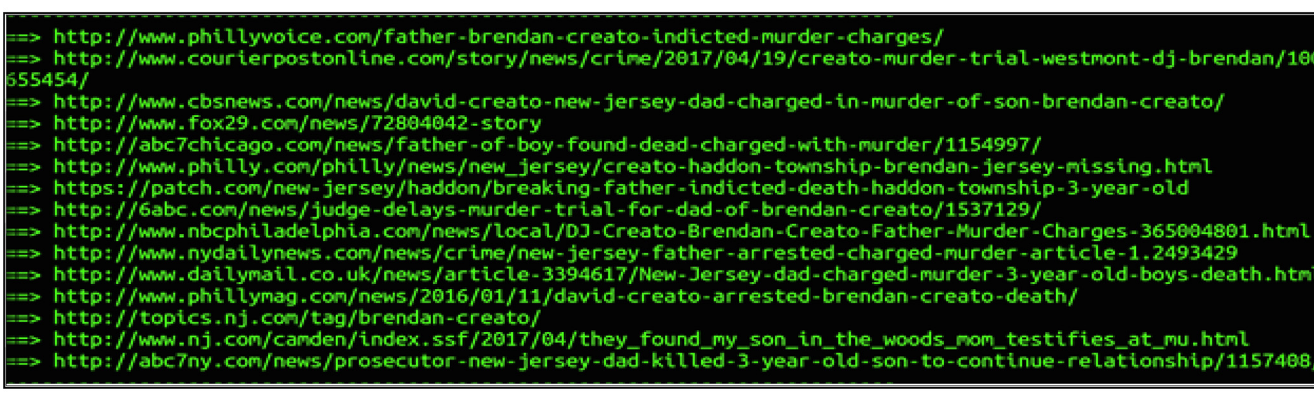
Table 9
Comparison of the proposed appraoch on state-of-the-art datasets.

Method	Name of dataset	Input type	Classifier	Acc (%)	F1 (%)	P (%)	R (%)
Zubiaga et al. 2016)	PHEME	Text	CRF	–	60.7	66.7	55.6
Shu et al. (2017)	FakenewsNet:	Text	SVM	86.4	87.0	84.9	89.3
	(1) BuzzFeed			87.8	88.0	86.7	89.3
	PolitiFact						
Horne and Adali (2017)	BuzzFeed election	Text	SVM	77	–	–	–
Proposed Method	PHEME	Text	Rule-based	–	**69.3**	**73.2**	**65.8**
	(1) BuzzFeed	Text	Rule-based	**85.3**	**86.77**	**85.2**	**88.4**
	(2) PolitiFact			**88.0**	**88.34**	**87.9**	**88.8**
	BuzzFeed election	Text	Rule-based	**86**	–	–	–

(a)



(b)



(c)

Fig. 6. Results on the real news (a) Input Image (b) Google search (c) Top 15 Google search.

5. Results

In order to test the performance of the proposed algorithm, an online test has been conducted on fake news. The sample output on real news and fake news are as shown in Figs. 6 and 7. Table 8 shows how reality parameter reacts on the local, national, and international news. The output results are analyzed in terms of the reality parameter (Rp), which is set at 40%, and the number of links is set to 15, giving the highest accuracy. It was observed that the match between the summary of the content of the web page with the search query was ranging from 0 to 70%, whereas the summary was ranging from 40 to 65 words each. The proposed system outperformed the state-of-the-art system giving an accuracy of 85%. The experimental results show that the proposed system does not perform well on a local/regional news event.

Table 10 gives the variance in the value of reality parameter with the type of news. We observe that the system fails to classify the local news, as the value of Rp is very low for

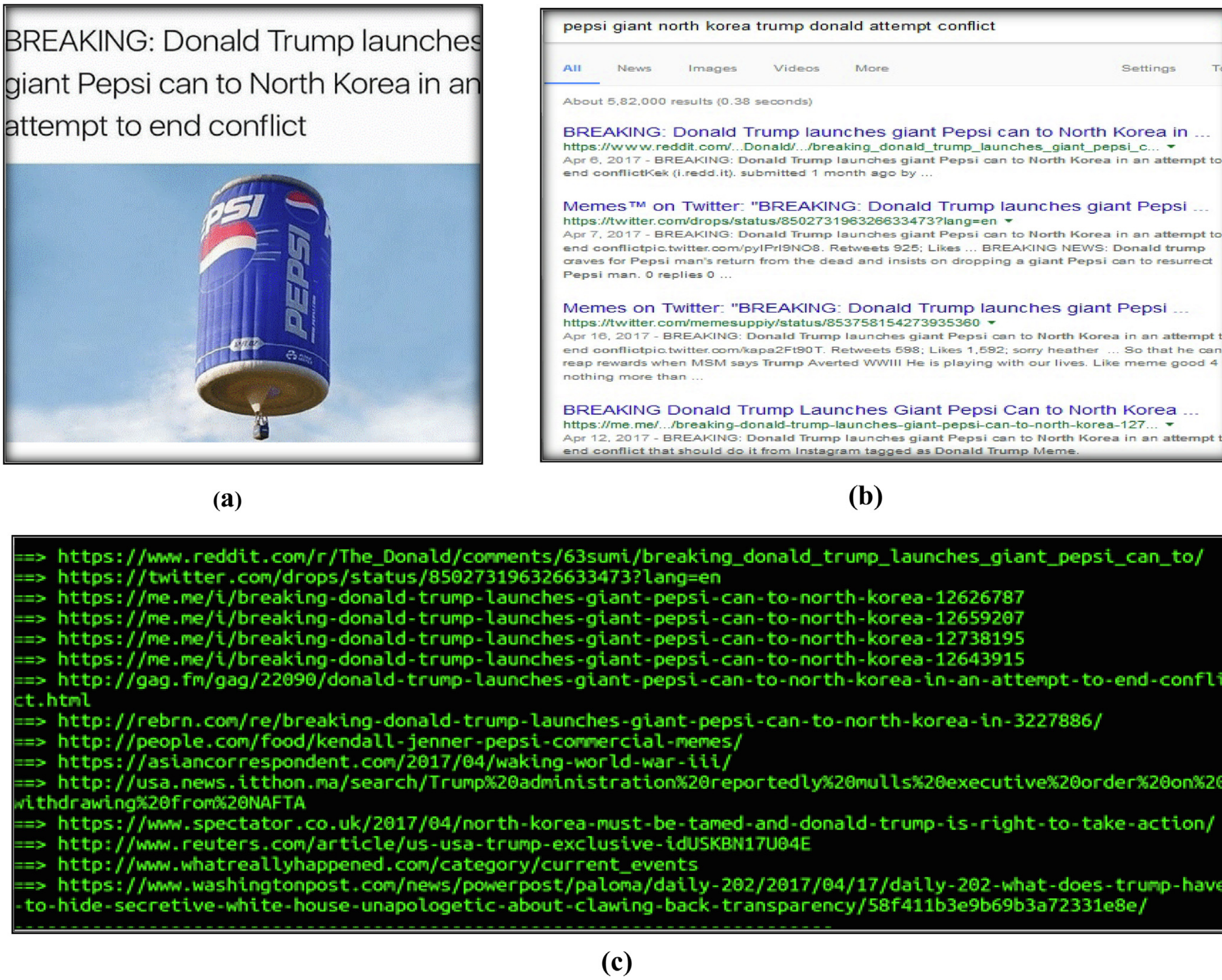


Fig. 7. Results of a fake image after going through our system (a) Input Image (b) Google search (c) Top 15 Google search.

Table 10
Reality parameter vs. news type.

Type of News	Local news	National news	International news
Reality Parameter (Rp)	0.0–0.20	0.26–0.80	0.46–0.86

it. For national news, the value of Rp is quite effective (some false negatives will occur), whereas for international news, the value is very effective.

6. Conclusion and future work

In this paper, we have developed a novel algorithm which can detect fake news events. Reality parameter performed best when its value is 40%, which give us 85% accuracy with the number of links being 15. Furthermore, the match between a summary of content and the search query seems to range between 0 and 48%. During our experimentation, a problem was faced during text extraction from images as for some images; we were not able to extract the text correctly because of various image characteristics like text with shadowing effect. The proposed system addresses the fake news problem for both national and international news. The system seems to fail to classify local news, as the news does not get enough heat for major players to cover them.

Future work can be based on the improvement in the process of entity extraction for images text if the image is having a large amount of text, as this will directly affect the Google search results. Moreover, integration of various social media handler of credible media houses or newspapers for authentication of a news event with the current system might further improve the accuracy.

Declaration of Competing Interest

There is no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cogsys.2019.07.004.

References

Ahmet, A., Derczynski, L., & Kalina, B. (2017). Simple open stance classification for rumour analysis. arXiv preprint arXiv: 1708.05286.

Aiello, L. M., Petkos, G., Martin, C., Corney, D., Papadopoulos, S., Skraba, R., ... Jaimes, A. (2013). Sensing trending topics in Twitter. *IEEE Transactions on Multimedia, 15*, 1268–1282.

AlRubaian, M., Al-Qurishi, M., Al-Rakhami, M., Rahman, S. M. M., & Alamri, A. (2015). A multistage credibility analysis model for microblogs. In *2015 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)* (pp. 1434–1440).

Alrubaian, M., Al-Qurishi, M., Hassan, M. M., & Alamri, A. (2018). A credibility analysis system for assessing information on twitter. *IEEE Transactions on Dependable and Secure Computing*, 661–674.

Boididou, C., Papadopoulos, S., Zampoglou, M., Apostolidis, L., Papadopoulou, O., & Kompatsiaris, Y. (2018). Detection and visualization of misleading content on twitter. *International Journal of Multimedia Information Retrieval, 7*(1), 71–86.

Bouazizi, M., & Ohtsuki, T. O. (2016). A pattern-based approach for sarcasm detection on twitter. *IEEE Access, 4*, 5477–5488.

Byungkyu, K., O'Donovan, J., & Höllerer, T. (2012). Modeling topic specific credibility on twitter. *Proceedings of the 2012 ACM international conference on intelligent user interfaces*.

Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on Twitter. *Proceedings of the 20th international conference on world wide web, Hyderabad, India*.

Chang, C., Zhang, Y., Szabo, C., & Sheng, Q. (2016). Extreme user and political rumor detection on Twitter. *Advanced data mining and applications. ADMA 2016*.

Cheung, M., She, J., & Jie, Z. (2015). Connection discovery using big data of user-shared images in social media. *IEEE Transactions on Multimedia, 17*, 1417–1428.

Chidiac, N. M., Damien, P., & Yaacoub, C. (2016). A robust algorithm for text extraction from images. *39th international conference on telecommunications and signal processing (TSP), Vienna*.

Dadkhah, S., Manaf, A. A., Hori, Y., Hassanien, A. E., & Sadeghi, S. (2014). An effective SVD-based image tampering detection and self-recovery using active watermarking. *Signal Processing: Image Communication, 29*(10), 1197–1210.

Elkasrawi, S., Dengel, A., Abdelsamad, A., & Bukhari, S. S. (2016). What you see is what you get? Automatic image verification for online news content. *12th IAPR workshop on document analysis systems (DAS), Santorini*.

Ferreira, A., Felipussi, S. C., Alfaro, C., Fonseca, P., Vargas-Muñoz, J. E., Dos Santos, J. A., & Rocha, A. (2016). Behavior knowledge space-based fusion for copy–move forgery detection. *IEEE Transactions on Image Processing, 25*(10).

Hoffman, P. (2013). Scrapy | a fast and powerful scraping and web crawling framework. Scrapy. [Online]. Available: https://scrapy.org/.

Holscher, E. (2015). Selenium with python. [Online]. Available: http://selenium-python.readthedocs.io/.

Horne, B. D., & Adali, S. (2017). This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. *Eleventh international AAAI conference on web and social media, Canada*.

Jin, Z., Cao, J., Jiang, Y. G., & Zhang, Y. (2014). News credibility evaluation on microblog with a hierarchical propagation model. *IEEE international conference on data mining, Shenzhen*.

Jin, X., Cao, J., Zhang, Y., Zhou, J., & Tian, Q. (2017). Novel visual and statistical image features for microblogs news verification. *IEEE Transactions on Multimedia, 19*(3), 598–608.

Li, Y., Dai, W., Ming, Z., & Qiu, M. (2016). Privacy protection for preventing data over-collection in smart city. *IEEE Transactions on Computers, 65*, 5.

Liu, X., Nourbakhsh, A., Li, Q., Fang, R., & Shah, S. (2015). Real-time rumor debunking on twitter. *Proceedings of the 24th ACM international on conference on information and knowledge management, New York, NY, USA*.

Niu, X. (2008). On two major issues in the current network news communication. *IEEE international symposium on IT in medicine and education, Xiamen*.

Pang, J., Jia, F., Zhang, C., Zhang, W., Huang, Q., & Yin, B. (2015). Unsupervised Web topic detection using a ranked clustering-like pattern across similarity cascades. *IEEE Transactions on Multimedia, 17*(6), 843–853.

Pasquini, C., Brunetta, C., Vinci, A. F., Conotter, V., & Boato, G. (2015). Towards the verification of image integrity in online news. *2015 IEEE international conference on multimedia & expo workshops (ICMEW), Turin*.

Qin, Y., Wurzer, D., Lavrenko, V., & Tang, C. (2016). Spotting rumors via novelty detection. *CoRR*, vol. abs/1611.06322.

Rashed, K. A., Renzel, D., & Klamma, R. (2011). Trust-aware media quality profiles in fake multimedia detection. *2011 Workshop on multimedia on the web, Graz*.

Shu, K., Wang, S., & Liu, H. (2017). Exploiting tri-relationship for fake news detection. arXiv: 1712.07709v1.

Steven, B., Loper, E., & Klein, E. (2009). Natural language toolkit. *Natural language processing with Python*. O'Reilly Media Inc.

Tang, S., Blenn, N., Doerr, C., & Mieghem, P. V. (2011). Digging in the Digg social news website. *IEEE Transactions on Multimedia, 13*, 1163–1175.

Tang, M., Mao, X., Guessoum, Z., & Zhou, H. (2013). Rumor diffusion in an interests-based dynamic social network. *The Scientific World Journal*, 1–10.

Vosoughi, S. (2015). Automatic detection and verification of rumors on Twitter Doctoral dissertation. Massachusetts Institute of Technology.

Vosoughi, S., Mohsenvand, M. N., & Roy, D. (2017). Rumor Gauge: Predicting the veracity of rumors on twitter. *ACM Transactions on Knowledge Discovery from Data (TKDD), 11*(4).

Wessel, M., Thies, F., & Benlian, A. (2016). The emergence and effects of fake social information: Evidence. *Decision Support Systems, 90*, 75–85.

Yang, F., Liu, Y., Yu, X., & Yang, M. (2012). Automatic detection of rumor on Sina Weibo. *Proceedings of the ACM SIGKDD workshop on mining data semantics, New York, NY, USA*.

Zampoglou, M., Papadopoulos, S., & Kompatsiaris, Y. (2017). Large-scale evaluation of splicing localization algorithms. *Multimedia Tools Appl., 76*(4), 4801–4834.

Zhang, X., & Ghorbani, A. A. (2019). An overview of online fake news: Characterization, detection, and discussion. *Information Processing and Management*.

Zhang, Z., Zhang, Z., & Li, H. (2015). Predictors of the authenticity of Internet health rumours. *Health Information & Libraries Journal, 32*(3), 195–205.

Zhu, C., & Wu, G. (2011). Research and analysis of search engine optimization factors based on reverse engineering. *Third international conference on multimedia information networking and security, Shanghai.*

Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys, 51*, 32:1–32: 36.

Zubiaga, A., Liakata, M., & Procter, R. (2016). Learning reporting dynamics during breaking news for rumour detection in social media. arXiv: 1610.07363.