International Journal of Data Science and Analytics (2025) 20:4149–4176
https://doi.org/10.1007/s41060-025-00727-w

REVIEW

A review of web infodemic analysis and detection trends across multi-modalities using deep neural network

Priyanka Meel[1] · Chahat Raj[1] · Bhawna[1]

Received: 20 September 2024 / Accepted: 24 January 2025 / Published online: 5 February 2025

Abstract

The proliferation of disinformation and misinformation across diverse digital platforms poses a significant societal challenge. Previous work in this area adequately addresses the false news detection on the online text shared. The technological platforms that have enough study and research carried out on them include the web news, as well as Twitter. This emerging field is gaining grounds for Facebook, Reddit, WhatsApp, YouTube, among other social media apps. Online data are analyzed using different modalities of text, images, videos, and speech with other sources of influence. It ended with the effectiveness of multi-modal integration in the identification of misinformation. To combat this issue, the development of robust and accurate detection techniques is imperative. This review delves into the multifaceted nature of this problem, exploring the intricacies of multi-modal fake news detection. It examines how the integration of text, image, video, and audio modalities can enhance detection accuracy. At present, there is an abundance of surveys consolidating textual fake news detection algorithms. This review primarily deals with multi-modal fake news detection techniques that include images, videos, and their combinations with text. We provide a comprehensive literature survey of eighty articles presenting state-of-the-art detection techniques, thereby identifying research gaps and building a pathway for researchers to further advance this domain.

Keywords Fake News Detection · Visual Data · Multi-modal · Deep Learning · Neural Networks

1 Introduction

An influx of online data has gone hand in hand with dramatically increasing diffusion of misinformation or what is often called the "infodemic." Such false, partly wrong narratives can make great social relevance, such as persuading public opinion or fueling violence. Reliable, efficient methods for detecting and analyzing such harmful content would be crucial to helping dampen the influence of this dangerous content [1]. A review, which would provide an inclusive overview of the latest contributions across this subject, would intend to analyze different kinds of modalities for the spreading of misinformation through texts images, video, audio. Current advanced methods are pushing researchers really well forward with the development of highly complex models that, through deep neural networks, could identify and interpret infodemic content in multiple modalities [2] [3] The article presents recent methods, challenges, and directions in the analysis and detection of a web-based infodemic. Let's remove the subtleties of each modality [4] Toward this end, all the advantages and disadvantages of several deep learning architectures offered by Transformers, Recurrent Neural Networks, and CNNs have been covered. Multimodality plays an important role in a manner that it enhances one's detection accuracy by covering rich information contexts.

WhatsApp is already famous for spreading false news and all that sort of thing on the platform. This time, a message containing misinformation about WhatsApp has been found spreading on a messaging app. According to this message, changes have been brought into being regarding how WhatsApp work. The message reads, "Two blue ticks and one red tick means the government can take action against while three red ticks will mean that the government has started court proceedings against you." The information contained in the

✉ Bhawna
bhawna1021@gmail.com

Priyanka Meel
priyankameel86@gmail.com

Chahat Raj
chahatraj58@gmail.com

1 Department of Information Technology, Delhi Technological University, New Delhi 110042, India

Fig. 1 **a** Visual Fake news of three red ticks in WhatsApp application, [5] (COURTESY: PIBFact Check) **b** Digitally morphed image of a child with three eyes, [6] (COURTESY: Twitter) **c** Digitally altered fake news or image of Mark Zuckerberg with Prime Minister Narendra Modi sitting alongside wearing Bhagwakurta [7] d) Pak Ambassador showing fake pic at UNGA, [8] (COURTESY: BBC)

message is completely fake [5] (Fig. 1a). A similar three red tick message had also gone viral last year and was debunked.

With the advent of many digital tools, it is easier to spread fake news via Facebook, Twitter and other related platforms. Figure 1b shows one of the examples of digitally modified fake news via image is digitally morphed fake news of a child with three eyes [6]. Figure 1c shows digitally altered fake news or image of Mark Zuckerberg with Prime Minister Narendra Modi sitting alongside wearing Bhagwakurta [7].

One of the most famous examples of spreading fake news on international media was, when Pakistani Permanent Representative to the United Nation Maleeha Lodhi held up the photograph [8], claiming that, it showed a Kashmiri girl, whose face bore the marks of Indian palette gun wounds. After that, India blasted Pakistani representative for using an image of a Gazan girl at the United Nation Forum to illustrate Indian violence in the Kashmir in Fig. 1d. Moreover, the photo or the image was taken during Survivor of 2014 war in Gaza and shared 17 years old, "Ravuya Abu Jomaa," who was an alleged victim of an Israeli attack.

It leaves with a great impression very rapidly. It is a fertile ground that allows changes and modifications and even forgeries. Advanced technology allows plenty of applications and methods of editing. It will change the images and videos so that information consumers are deceived. Online social media platforms allow users to contribute their data regardless of its authenticity. Recent sightings included a picture of a shark on a Houston freeway, which made national headlines during Hurricane Harvey. That picture was widely shared and retweeted to make people very fearful. Misinformation has been so outrageous that even US President Donald Trump could not escape from it and went on to retweet a fake video that stated anti-malaria drugs could cure coronavirus, which was later brought down when proven fake [9] [10].

1.1 Motivation

A huge amount of literature and resources are reported on text-based detection about the study of trends in the identification of false information. Since its origination, many frameworks and mechanisms for detection based on the use of textual features have developed after this problem area. Mainly due to their popularity, machine learning and deep learning algorithms are applied extensively within this field for offering solutions.

- The increasing prevalence: The increased number and complexity of fake news spread through writings, images, and audio-visual media.
- Requirement of an Inclusive Detection System: No general framework is available that covers all types of misinformation.
- A primary importance of multimodal analysis is the ability to incorporate textual and visual indicators for better identification.
- Identification of Gaps in Research: The need to address new challenges and potential avenues in the subject.

With the increasing trend of false news coming through text, image, and video calls, an effective detection mechanism is quite in demand. Great strides have been achieved, however, in areas concerning multimodal analysis, deepfake detection, and real-time detection. Advanced techniques combined with interdisciplinary collaboration can check the spread of false information and therefore protect information integrity.

2 Research Question

The spread of false information and fake news is among the most acute challenges in the modern digital environment. Social media, and quite a few other services on the Internet, become channels for free and uncontrolled spreading of misinformation, which can result in disastrous consequences: from the change of opinion of public masses to damage to democratic values and the further drawing out of citizens' confidence in institutions. Measures should be developed to effectively detect and counter fake news.

- How well do the machine learning and deep learning performs in multi-modal fake news detection tasks?
- What are the trends in fake news detection and data manipulation detection methods?
- What are the trends across fake news modalities and benchmark datasets?

From these research questions' resolutions, we will assist in the development of effective tools and strategies that work toward the eradication of the spread of fake news and toward a more enlightened, informed public.

2.1 Contribution and Organization of the paper

To the best of our knowledge, this is the first data modality-based review in fake news detection. The contributions of this study are as follows:

- Analyzing and identifying the techniques utilized in multimodal fake news detection tasks.
- Comparing these techniques based on their applications, advantages, and disadvantages.
- Providing a comprehensive review of remarkable work done in the domain, discussing popular techniques, datasets used, and results obtained.
- Providing a detailed summary of multi-modal usable datasets for fake news detection.
- Comparing the efficiencies of available literature and their work in terms of evaluation parameters utilized.
- Identifying the research gaps in multimodal fake news detection methods and enlisting potential research directions.

The organization of this paper is as follows: Sect. 1 discusses the current Infodemic situation and introduces readers to the problem domain. Section 2 provides a detailed review of the literature and important architectures built for fake news detection tasks. In Sect. 3, we described that how we have done the data collection & analysis according to the existing challenges in multimodal fake news detection. Section 4 discusses several detection methods used so far across textual and visual modalities. Section 5 provides fake news detection mechanisms. Section 6 contains the trending datasets and performance comparison based on the novelty of the architecture. In Sect. 7, we discuss potential future research directions. Section 8 concludes this review by summarizing our work and imparting motivation and future research directions to readers.

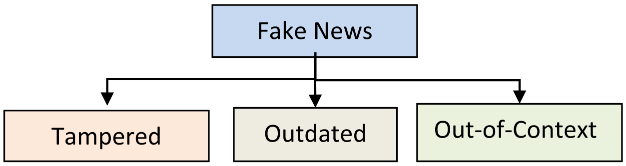

Fig. 2 Types of fake images

3 The Canopy of Fake News Detection

Image tampering has become easier than ever, given the advances in photo editing tools. It is crucial to detect such forgeries to keep a check on fake news data [11] Fake images accompanied by fake news are categorized as in Fig. 2. The categories are tampered/edited images, outdated images used with a later situation, and images out-of-context with the accompanying text. Mainly due to their popularity, machine learning and deep learning algorithms are applied extensively within this field for offering solutions. These include sentiment analysis, text mining, stance classification, similarity analysis, etc. [12] Texts are analyzed based on their sentence structures, words, punctuations, tone, grammar, and pragmatics Textual fake news analysis is the domain well explored and worked upon by many researchers.

Elkasrawi et al. [12] allowed users to verify images present online using metadata and feature analysis automatically. The approach has two phases: first involves checking the image for its authenticity, whereas, in second, it is verified if the image has any alterations. The first phase attempts to find other occurrences of an image on the web using Google Image Search. This results in the appearance of the query image in its similar or manipulated versions. Along with several versions of the image, metadata like URL of the image, article in which it appears, publishing date, time, thumbnail, etc., are retrieved. The image's authenticity is validated by matching the timestamps of the query image and search result image. If the image dates back from the date of the query image, it is regarded as fake.

In Fig. 3a, there are two images original image showing only 1 bird and the tempered image showing 2 birds. Figure 3b again showing 2 images first one with a face of Mr. Donald Trump that is imposed on the original image represented in second part. In several occurrences, the algorithm uses k-means clustering for the resulting images and their timestamps. In the second phase, image matches are retrieved, and alteration detection is performed. This is done using edge comparison and checking image alignment. The algorithmic features are conglomerated and deployed in a Chrome browser extension for image-based fake news verification. Approaches utilized to detect these fakes include supervised deep learning techniques that require huge training data.

3.1 Current approaches in fake news detection

Zhang et al. [1] worked on the difficulty of preventing false information in a cyber-physical realm, highlighting the short courage's of current strategic because of delayed semantic modeling. To address this, a convolution- based rural framework for effective feature exuviation is combined with Chinese characters as processing units in a Chinese specific deep learning model. Gramigna et al. [4] in a 2021 study published in the ASEAN Journal for Public Opinion Research that scanned 419 instances of social media fake news, found that online media, not traditional sources like print and broadcast news, accounted for the vast majority of misinformation. Health, religion, and politics dominated the trends in the spread of fake news. A 2021 study by AJPOR found that Facebook and Twitter stood out as the leading two media through which the wrong information circulated. From the 90 instances of false information analyzed, 81 percent were found on Twitter. Moreover, half of the misleading reports were recorded in video form, while about one-third consisted of images and text. The spread of fake news was especially rapid during COVID-19 pandemic. Zhang et al. [2] acknowledges the social impact of tools such as "deep fakes" and "shallow fakes" by focusing on the localization of modified areas for forged photos with a publicly available dataset and code the research presents an accurate method for detecting modified regions in shallow and deep fake photos. Wang et al. [13] built a neural network framework for incoming real-time events. This Event Adversarial Neural Network (EANN) framework can handle event-invariant features, thus allowing the detection of fake news on freshly arriving events. With three components in the framework, the first component is a multi-modal feature extractor. Text features are extracted using TextCNN, and a pre-trained VGG-19 is used for visual feature extraction after fine-tuning the hyperparameters [14]. The second component is the fake news classifier, which is built upon the multi-modal feature extraction layers [15] [16] [17] Classification is performed using a softmax layer for predictions. The third component is an event discriminator that classifies a post into one of the K events. Its application is to remove event-specific features from the posts and capture only invariant features across all events. EANN is capable of classifying fake news incoming from any type of event.

All other image classification tasks that use deep neural networks utilize CNN models, which have been pre-trained on the ImageNet dataset. Jin et al. [18] highlight that to use CNNs for fake news classification tasks, it is required to train the models on a specific fake news dataset. For image credibility analysis of online news on visual data, they collect an auxiliary dataset of fake images from tweets with 328 K rumor images. These images are labeled in terms of their credibility polarity. A collection of 0.6 M images is used

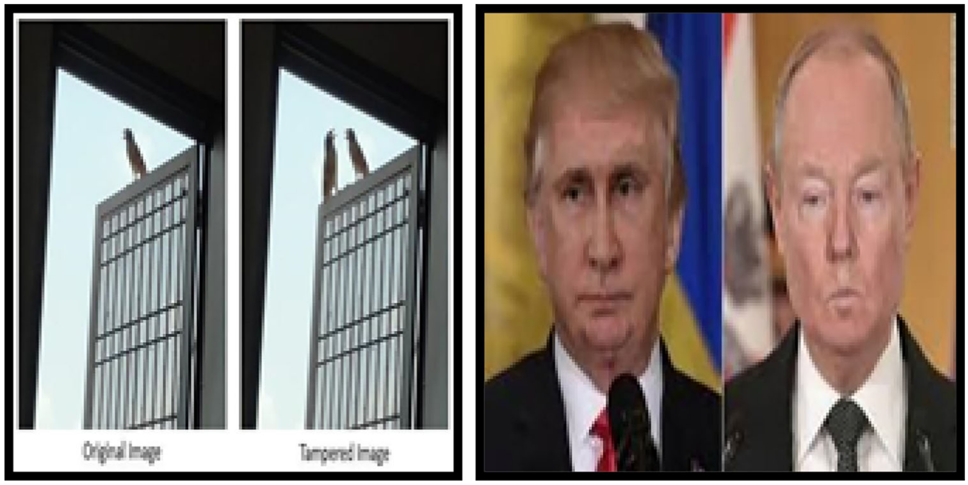

Fig. 3 **a** Original vs. tampered image, **b** Face-swapped images

to pre-train a convolutional network with architecture similar to AlexNet. The model is trained using iterative transfer learning. This domain transferred learning algorithm provides favorable results for fake news classification.

[12] Prominent example of spreading fake news through audio overlay on video is when fake news surfaced on international media. Fake BBC clip on Ukrainian politician selling arms to the Hamas. The BBC did not publish the video claiming that a Ukrainian politician sold weapons to Hamas. Text overlaid on the video at the start of the clip reads, “A Ukrainian politician may be involved in arms sales to Hamas." Reference video has been deleted from the media platforms. [13] One more incident wherein a fake post claims to show, “Live encounter in Kashmir”. After analyzing and checking the facts, it was crystal clear that the video was old and shows a Joint Military Exercise by India and USA, which had been conducted in 2016 in a village called Chaubattia in Utrakhand. Text, images, and videos compose a major actuating portion of the information on the internet and hence blot the gray areas of fake news. Since fake news and its critical nature came into the picture, several scientific developments have been made in textual fake news detection. Fake news detection technologies using NLP, text classification, vector-space models, rhetoric structure theory, opinion mining, sentiment analysis, graph theory, deep neural networks, and others have been created, reviewed, and summarized by fellow researchers [16] [17] [18] Probing of image fake news is comparatively lower, and of videos, it is negligible. Most of the information users encounter on the internet is accompanied by visual representations, either using an image, video, or other modalities.

Qi et al. [19] are the first to use multi-domain information to classify fake and real images proposing MVNN (Multi-domain Visual Neural Network). They expressed that fake news images are constructed of different physical and semantic features than in real news images. They fused the frequency domain and pixel domain feature extraction subnetworks using an attention mechanism to identify real and fake images. Their proposed network consists of three components: a CNN-based frequency domain sub-network, a pixel domain sub-network built with CNN-RNN to extract semantic features, and a fusion sub-network. As a whole, their network classifies fake news by using only image features, disregarding linguistics, which is a limitation of the task.

Vishwakarma et al. [20] performed FND by scrapping and authenticating web searches. The work proposes an architecture that enables the verification of text-embedded images. Extraction of text from images is supported by Optical Character Recognition (OCR). The extracted text undergoes Name Entity Recognition steps to obtain named entities mentioned in the text. These strings are used as queries for Google Search to find matching results. The resulting links are categorized into reliable or unreliable. Next, the entities extracted from images are checked against the search result articles' titles, further classifying them into real or fake. Pasquini et al. [21] also verified the integrity of online news. They have also demonstrated that fake news detection in visual has a dependency on visual forensics. Their focus is to detect news web pages that contain misleading images. Their framework automatically detects related news articles on the internet with similar images of an event. The detection is performed using metadata features present along with the images and by comparing similar features in the set of related images.

Using image forensics techniques for fake news classification, Huh et al. [22] utilized self-learned consistency to detect image manipulation and splicing. It is done by comparing patches of images. Tampered images show low consistency scores between patches. Models are evaluated for splice detection and splice localization on unannotated images. It is an efficient way to detect fake/manipulated images where copy-move forgery, object addition, or removal have been done. SAME (Sentiment-Aware Multi-modal Embedding) is a deep end-to-end embedding network that exploits users' sentiments to classify fake news. Cui et al. [23] have incorporated users' sentiments with images, profiles, content, and comments into a multi-modal framework. The model intakes post content, image, and user profile and feeds them to a Multi-Layer Perceptron (MLP), pre-trained VGG-19 network, and another MLP, respectively. The three modalities are fused using an adversarial mechanism. Jin et al. [24] performed a classification between real and fake tweets using a two-level classification model: topic level and message level. They proposed that tweets among themselves have a strong relationship in terms of event or topic, and tweets clustered in a topic have similar credibility values. It provided a better result than uni-level classification. Tweets with similar images (verifying by features such as resolution, image popularity, etc.) were clustered under the same topic. They extracted topic-level features and message-level content features, user features, and other available features and classified them on both levels. Deep neural networks have been successful in the classification of manipulated images from the originals. Table 1 summarizes all the tasks related to fake news detection that involve visual modalities in ascending order of their accuracy and F1-score. It helps to understand the necessary details of the related works easily.

From 2013 to 2024, Table 1 shows that GoogleNet and ResNet with DCT and HSV gives lowest F1-Score of 40% for Video annotations. When contextual cues have been used for video verification, the model gives 90% F1-Score.this means contextual embedding with CNN models is giving better results but Two-level classification of the model for text and images gives 5highest F1-Score of 94%. Meta-analysis for image verification gives 88%. Machine learning algorithm like naïve bayes, decision tree used for image classification gives highest accuracy of 97%. We have concluded that although Deep learning algorithms providing better F1-score, more refining and combing of contextual embedding can improve the results for accuracy also.

Topic-level classification results were fused into message-level feature-vector as an extra feature and then trained the classifier. Each tweet was given a separate label instead of labeling each event. Shu et al. [38] presented a way to utilize user profile features for fake news detection. They extracted and studied explicit and implicit user profile features, also studying which users were most likely to share real and fake news. They have studied users' geolocations, profile images (using CNN, pre-trained VGG16 model), and political bias. PCA was used for dimensionality reduction of profile features. They compared fake news detection performance using UPF (User Profile Features) to multiple approaches, including RST (Rhetorical Structure Theory), LIWC (Linguistic Inquiry and Word Count), RST_UPF, and LIWC_UPF concluding that UPF and UPF-allied techniques provided higher accuracies than others. They also proved that implicit and explicit features, when combined, provided greater results than each being used individually.

Ajao et al. [48] applied hybrid CNN and LSTM-RNN models for text and image classification. LSTM RNN was used to process and classify text sequences. Another model used was LSTM along with dropout regularization 0.2 to remove over-fitting. The third model incorporated a CNN layer after the word-embedding layer of the LSTM model. The models plain vanilla LSTM, LSTM-CNN model, and LSTM with dropout regularization performed in the order of decreasing accuracies 82%, 80%, and 74%, respectively. Under-fitting and lack of sufficient training data account for the low accuracy of the LSTM-drop model. They also showed that hybrid deep learning models' usage provides considerably good accuracy without requiring huge training data. Singhal et al. [49] have designed a model named SpotFake for detecting multi-modal fake news without any subtasks like finding correlations between textual and visual data. The model consists of a textual feature extractor module that uses BERT (Bidirectional Encoder Representations from Transformers), a visual feature extractor module that uses VGG19, and a fusion module using simple concatenation. The results outperformed state-of-the-art methods EANN and MVAE and provided higher accuracies equal to 77.77% and 89.23% on Twitter and Weibo, respectively.

Trumper [50] developed a web tool, 'Fake Tweet Buster,' for users to check a tweet's credibility. The user needs to enter a tweet URL, and the application provides a result as fake or legitimate. The tool works on reverse image search (Google Images and TinEye), user analysis, and crowdsourcing. The tool provides the user with matching images, image data (old or new), tweet information, and classification result. It allows the tool-user to enter his opinion about the tweet as fake, legitimate, or unsure. This crowdsourced data can be used in the future to provide value to credit score. Gupta et al. [25] extracted tweet information and images from Twitter related to Hurricane 'Sandy'. They analyzed those retweets contained many fake images (86%) rather than original tweets. The approach used temporal analysis to study the propagation of fake images over the Twitter network considered a graph with multiple nodes. They created graph networks for followers and retweets of each user studied. For classification, they used user-level features (giving poor results) and tweet level features (providing effective classification), experimenting with two machine learning techniques: Naïve Bayes classifier and J48 Decision Tree classifier, providing 91% and 97% accuracy, respectively. Lago et al. [37] created a fusion of image forensics algorithms to distinguish between real and fake images. The various approaches used are Image Forensics Methods (Error Level Analysis, Block Artifact Grid Detection, Double Quantization Likelihood Map, Median-filter noise residue inconsistencies detection, JPEG Ghosts,

Table 1 Summary of crucial work based on different methods & techniques

Year [Reference]	Method/ Technique	Task	Modality	Model Classification	Dataset	Accuracy/F1-Score
2013 [25]	Temporal analysis, Naive Bayes, decision tree	Fake image classification of Hurricane Sandy	Text, Image	Multi-model	Twitter	97% (Accuracy)
2015 [24]	Two-level classification (Topic and message-level)	Tweet labeling for FND	Text, Image	Binary	Twitter (Mediaeval 2015[1])	94% (F-score)
2016 [26]	CNN	Fake news detection	Text, Image, Source	Mono	Twitter	82.47% (F-score)
2016 [18]	Image splice detection, splice localization	Fake news detection	Image	Mono	Columbia, Carvalho, Realistic Tampering, In The Wild, Hays and Efros	91% (F- score)
2017 [27]	CNN, Iterative Transfer Learning	Fake news detection	Image	Binary	Weibo	77% (Accuracy)
2017 [28]	Visual and statistical features	Microblogs news verification	Image	Multi-Model	Sina Weibo	83.6% (Accuracy)
2017 [29]	CNN (frequency domain), CNN-RNN (pixel domain), Bi-GRU	Fake news detection	Image	Multi-Model	Twitter, Weibo	84.6% (Accuracy)
2017 [30]	CNN with Contextual Cues	Video verification	Video	Mono	IVC, FVC	90% (F-score)
2017 [31]	LSTM, att-RNN, VGG19	Rumor detection on microblogs	Text, Image	Multi-Model	Twitter, Weibo	78% for Twitter (Accuracy), 68% for Weibo (Accuracy)
2017 [32]	Text-CNN, Visual, Similarity features	Fake News Detection	Text, Image	Binary	PolitiFact, GossipCop	87.4% for PolitiFact (Accuracy), 83.8% for GossipCop (Accuracy)
2017 [33]	MAE, Bi-DNN, VSM	Semantic Integrity Assessment	Text, Image	Multi-Model	MAIM, Flickr30, MS COCO	(F-scores) 75% (MAIM), 89% (Flickr30), 94% (MS COCO)
2023 [34]	CNN	Fake news detection	Text, Image	Binary	TI-CNN	92% (F-score)
2018 [35]	CNN	News consistency verification	Text, Image, Location, Events	Binary	TamperedNews, News400	94% (Accuracy)
2018 [36]	VGGNet, sentiment analysis	Fake news detection	Text, Image	Multi-Model	PolitiFact, GossipCop	~ 75%, ~ 80% (Macro, Micro-F1)
2019 [37]	Image forensics, visual & textual analysis	Image tampering detection, text-image coherence detection	Text, Image	Multi-Model	Mediaeval2016, BuzzFeedNews, CrawlerNews	> 75% (Accuracy)

Table 1 (continued)

Year [Reference]	Method/ Technique	Task	Modality	Model Classification	Dataset	Accuracy/F1-Score
2019 [38]	Utilizing user profile features	Fake news detection	Text, Image	Mono	FakeNewsNet	> 90% (F-score)
2019 [39]	Text-CNN, VGG19	Fake news detection	Text, Image	Binary	Twitter, Weibo	71.5% for Twitter (Accuracy), 82.7% for Weibo (Accuracy)
2019 [40]	Bi-LSTM, VGG19	Fake news detection	Text, Image	Binary	Twitter, Weibo	74.5% for Twitter (Accuracy), 82.4% for Weibo (Accuracy)
2019 [19]	Metadata, feature analysis	Image verification	Image	Mono	Twitter, Weibo	72.7% for Twitter (Accuracy), 88% for Weibo (Accuracy)
2019 [27]	BERT, VGG19	Fake news detection	Text, Image	Binary	Twitter, Weibo	77.77% for Twitter (Accuracy), 89.23% for Weibo (Accuracy)
2020 [41]	DCT, HSV, SURF, AOF, ResNet, GoogLeNet	Video annotation	Video	Multi-Model	Twitter	40% (F-score)
2019 [42]	Cosine Similarity, Embedding similarity	Fact-Checking on image-claim pair	Text, Image	Binary	Snopes, Reuters	80.1% (Accuracy)
2021 [43]	Machine Learning Algorithms	Fake news detection	Text, Image	Multi-Model	Kaggle Fake News Dataset 2017	85.25% (Accuracy)
2019 [44]	Montage detection (feature-based approach), SIFT, SURF	Fake News Detection	Image	Multi-Model	COCO, INRIA	> 90% (Accuracy)
2023 [45]	MFPN (BERT) with Fusion	Fake news detection	Text, Images	Binary	Weibo and twitter	83.3%
2024 [46]	LSTM with attention mechanism, VilBert	Fake news analysis	Text, image	Multi-model	Buzzfeed, Medieval twitter, CIC dataset	96.27%
2024 [47]	Reverse image search, User analysis, crowdsourcing	Fake tweet and its user identification	Text, Image	Multi-class (fake, legitimate, not sure)	Twitter	80% (Accuracy), 84% (F-score)

[a]https://github.com/MKLab-ITI/image-verification-corpus/tree/master/mediaeval2015
[a]https://github.com/MKLab-ITI/image-verification-corpus/tree/master/mediaeval2016
[a]https://github.com/MKLab-ITI/fake-video-corpus
[a]https://github.com/KaiDMML/FakeNewsNet
[a]https://github.com/entitize/fakeddit
[a]https://drive.google.com/file/d/0B3e3qZpPtccsMFo5bk9Ib3VCc2c/view

Color Filter Array), Splicebuster, and with the use of CNNs. They also verified the coherence of online news with their accompanying images. Texts were analyzed using TF-IDF or its higher version, STF-IDF. To make classifications, they chose Random Forest and Logistic Regression.

Yang et al. [34] used convolutional networks for both textual and visual fake news detection. They analyzed latent features and explicit features of text and image sub-branches and fused them using early fusion to classify news. Huckle and White [32] utilized blockchain and cryptography to trace the origination of fake content. Their focus is to determine where the fake information comes from rather than analyzing its structure and features. Their approach lies in determining the cryptographic hash functions of the text and associated image. The verification is made by matching the hashes of the versions of an image occurring at different places. Same images from different occurrences resulting in different hashes indicate that the image has been altered. This principle forms the base of their fake news detection mechanism. Krishnan and Chen [35] identified tweets containing fake news using data mining, statistical analysis, reverse image search, and cross-verification. They have divided their framework into two components: Core and Website. The Core module fetches tweets, extracts the feature set, performs classification, and returns predictions. The Website module is mainly responsible for crowdsourcing and collecting users' guesses about the post's credibility. It also returns the final decision about the post to the end-user. Classification is performed using the J48 decision tree classifier and Support Vector Machine (SVM). The crowd sourced data is stored in a database for future re-training of classification models and performance improvement.

The NewsVallum approach introduced by Armano et al. [51] uses text-image semantics for fake news detection. It focuses on multimodality using text and image features for classification with deep neural networks and reinforcement learning. It evaluates the credibility of news spread online on a daily or hourly basis. Zhou et al. [41] exploited the similarity between text and images to detect if a news article is true. The proposed model comprises three modules where the first one is a multimodal feature-extractor, the second module is a classifier, and the third module determines the cross-modal similarity between text and image in the post. They have used a text-based one-dimensional convolutional network for text-classification, and image features are extracted using VGG. Both the feature extractors are combined using a concatenation operation. In the third module, the semantic relation between the text and image pair is calculated using cosine-similarity. This helps in identifying if the modalities of a post express the same meaning. A post is classified as fake or real, depending on all of these features.

Chen et al. [52] argue the need for an automatic fake news detector tool for evaluating the integrity of any news online. Using multi-modal features, Budack et al. [53] measured the consistency between the modalities for fake news verification. The proposed work evaluates the coherence between text and image data in an unsupervised manner. Textual module extracts persons, entities, and locations using Named-Entity Recognition (NER). POS Tagging is applied, and subsequently, embeddings are calculated using fastText. For visual features, the ResNet model is used for verification. Persons, locations, events, and scene context verification is performed in the cross-modal entity verification process. This model applies to real-world news classification. Parikh et al. [54] performed the task on tweet text and images providing a web application. The UI allows users to upload a screengrab of a tweet from which the model extracts useful information like tweet text, image, username, timestamp, location, etc., and predict the authenticity of a tweet using these features. It has become easier to create fake scenarios in videos by replacing, removing, or adding objects, adding text, and swapping face in the recent technological eras. These manipulated videos in which a person's face is swapped by another's termed as 'deep-fakes.' Such videos have created nasty issues of defamation. Forensics-based detection methods majorly detect copy-move manipulations or tampered frames. Neural networks have been applied to detect morphed faces in visual data.

Nixon et al. [55] annotated videos as real or fake to verify the news. They analyzed the text of news stories and videos circulating online and used their metadata to fact-check and annotate the videos. In textual analysis, the author checks the stories for correctness, distinctiveness, homogeneity, and completeness and then groups them under clusters. The videos related to these news stories were then retrieved and annotated based upon fragment information. Bagade et al. [56] developed a fact-checking web and mobile application 'Kauwa-Kaate' for full-article verification incorporating text, images, and videos. Their proposed system provides a user-friendly interface to query and fact-check information as and when they encounter it. The algorithm scrapes news articles from fact-checked and trusted news sources available on the internet and maintains a repository in the backend. The verification is carried out by matching the query item with news articles in the repository. Devoting a platform entirely for fact-checking, is a very practical method for users to verify fake news. Using several tweet-based and user-based features, Boididou et al. [57] have introduced a model that predicts tweets as fake or real depending upon the majority vote from individual classifiers that use different features. The approach has displayed a successful classification of news from a variety of events. An event rumor detection mechanism for Sina Weibo has been designed by Sun et al. [58]. The model uses many features, which are content-based linguistic features, user-based features, and multimedia features. The model is suitable for detecting rumors in the form of text, image, and video.

3.2 Data manipulation detection techniques

Detecting fake news in visual data is closely related to identifying manipulations in images and videos. The utilization of data manipulation detection algorithms for fake image/video detection offers a worthy scope. In this sub-section, we identify some of the important manipulation detection techniques that could be useful in fake news detection tasks. A new convolutional layer has been proposed by Bayar and Stamm [59] to classify unaltered and manipulated images. Fake images can be of the misleading type where news content and accompanying image are unrelated or of the tampered type where images are forged to create a fake scenario. Detecting forged images can prove highly beneficial to detect fake news and news containing fake images. The new layer can learn manipulation detection features without needing to extract preliminary features. This layer has been incorporated into a CNN architecture to detect multiple manipulations. For model training, ReLU activation and SGD (Stochastic Gradient Descent) is used. CNNs train faster with ReLU. Binary classification classifies images into unaltered vs. tampered images with 99.31% accuracy. Multi-class classification is used to classify images based on types of forgeries used: median filtering, Gaussian blurring additive white Gaussian noise, resampling vs. authentic image with 99.10% accuracy.

Sabir et al. [60] and Jaiswal et al. [61] have detected image repurposing where unaltered images are put together with false metadata in a news item. Pomari et al. [62], Zampoglou et al. [63], and Wu et al. [64] detected fake images by checking if the images or their portions have been sliced. In Table 2, a summary of supportive works is provided, which utilize data forensics mechanisms to identify tampered visual data.

Identifying data manipulation helps maintain the integrity and reliability of the data. There are various methods that exist and can detect potential manipulation techniques. Table 2 represents the summery of the Examples of data manipulation detection techniques and is also distributed technologically. Different kinds of statistical analysis are also able to detect abnormalities in the distributions of data, such as outliers or inconsistencies, which may suggest manipulation. While it is not possible to get 100% accuracy, but VGG 19 showing the best results for generative adversarial networks for images. CNN has also shown 99.45% accuracy that is close to VGG19. With these techniques combined, robust systems intended to detect data manipulation events and data manipulation intents could be established, thereby ensuring data integrity for decision making. For data manipulation detection, deep learning algorithms are giving far better results but for getting better results for other evaluation metrics like maP, IoU etc. needs to be considered.

Pomari et al. [62] did so by making use of illumination inconsistencies in the image. Computer-generated images and videos have gained huge popularity in the present scenario. It has become easy to generate fabricated content through computers. Such content is entirely unreal or mixed with some kind of real-world entities. The result is a fake product that is not reliable at all. Tariq et al. [69], Marra et al. [67], Nguyen et al. [70], Rahmouni et al. [81] and Rezende et al. [82] have identified fake images generated by computer machines. Nguyen et al. [70] did this using Modular CNN. Sabir et al. [39], Zhang et al. [27], Zhou et al. [29], Dang et al. [83] have detected tampered faces in images, which can be utilized in detecting fake news where faces of celebrities have been swapped to create fake scenarios.

Wu et al. [84] have used supervised learning to trace various types of manipulations like copy-move forgery, object removal, splicing, and other unknown tampering in images. Photoshop is a widely used tool used by content creators to modify visual content. Wang et al. [85] have detected variations created in images using Adobe Photoshop. Wu et al. [36] created BusterNet for copy-move forgery detection. It is a type of forgery where an object from an image is removed from its original location and moved to a different place in the same image. Li et al. [75] used RGB color components in the images to detect changes that occurred due to tampering. Nataraj et al. [74] detected GAN-generated images using co-occurrence matrices. These matrices describe the distribution of co-occurring pixel values or colors. Steinebach et al. [44] recognized image montages for fake news detection. In image montaging, two or more images or their parts are arranged together by cutting, overlapping, pasting, etc., to make a composite image. Korshunov and Marcel [86] and Guera and Delp [72] addressed facial manipulations in videos where the face of a person is replaced by another. They used deep neural networks for the task. Guera and Delp [72] also contributed with a dataset of 300 deep-fake videos extracted from websites. They classified videos using CNN and LSTM into pristine and deep-fake categories. CNN has been used for feature extraction from video frames, concatenated and propagated to LSTM for analyzing sequences temporally. This architecture allows detecting fake videos as short as 2 s of length. Korshunov and Marcel [86] showed that using static frame features correspond to higher accuracies than using audio-visual analysis.

Nguyen et al. [73] identified forgeries like replay attacks, computer-generated images/videos by building a capsule network with CNN layers. Videos are analyzed at frame level, and the probabilities of fake and real of every frame are averaged to generate results for the video. Yang et al. [65] detected swapped faces in images and videos using CNN. Korshunov and Marcel [66] performed tampered video detection using

Table 2 Examples of data manipulation detection techniques

Year [Reference]	Technique	Task	Modality	Model Classification	Dataset	Accuracy/F1-Score
2014 [65]	CNN	Face Spoofing Detection	Image	Multi-Model	CASIA, REPLAY-ATTACK	< 5% (HTER)
2016 [59]	New convolutional layer	Image manipulation detection	Image	Binary	Various camera images	~ 99.10% (Accuracy)
2018 [66]	Speaker inconsistency detection	Tampered video detection	Video, Audio	Binary	VidTIMIT, AMI, GRID	< 1% (EER)
2018 [60]	CNN (VGG19), Word2Vec	Image repurposing detection	Text, Image, Location	Binary	MEIR	80% (F-score)
2018 [67]	Conventional methods, CNN (DenseNet, InceptionNetV3, XceptionNet)	Image translation detection on compressed and uncompressed images	Image	Multi-Model	CycleGAN	89% (Accuracy)
2018 [62]	CNN (ResNet50), Illumination Maps	Splice detection	Image	Binary	DSO, DSI, Columbia	96% (Accuracy)
2023 [68]	3D-CNN, CNN, openSMILE	Deception detection	Text, Video, Audio	Multi-Model	Courtroom trials	96.14% (Accuracy)
2018 [69]	CNN ensembles	Fake image detection	Image	Binary	CelebA, PGGAN	99.99% (AUROC)
2018 [70]	VGG19	GANs vs. real image detection	Image	Mono	RAISE, Rahmouni	99.9% (Accuracy)
2019 [71]	Mixed Adversarial Generators	Fake image detection	Image	Binary	FantasticReality	61% (precision)
2019 [39]	RNN, CNN (ResNet50, DenseNet)	Face manipulation detection	Video	Multi-Model	Deepfake, Face2Face, FaceSwap	96.9% (Accuracy)
2019 [72]	RNN, CNN (InceptionV3), LSTM	Fake video detection	Video	Multi-Model	Deepfake videos, HOHA	> 97% (Accuracy)
2019 [73]	Capsule Forensics, VGG19	Forged image and video detection	Image, Video	Binary	Deepfake	99.23% (Accuracy)
2019 [74]	Co-occurrence matrices, CNN	GANs Image detection	Image	Mono	cycleGAN, starGAN	99.45% (Accuracy), 93.42% (Accuracy)
2020 [75]	Color disparities (DCGAN, WGAN-GP, Pro- GAN, STYGAN, etc.)	Fake image detection	Image	Multi-Model	CelebA, LFW, generated images	> 90% (Accuracy)
2021 [76]	MCNN	Fake news detection	Text, Image, Video	Mono	MCG-FNeWS, Twitter, PolitiFact	94.7% (Acc) 94.6% (F1) MCG-FNeWS, 78.4% (Acc) 83.1% (F1) Twitter, 88.4% (Acc) 91.7% (F1) PolitiFact

Table 2 (continued)

Year [Reference]	Technique	Task	Modality	Model Classification	Dataset	Accuracy/F1-Score
2022 [77]	Transformer (Encoder/Decoder) + MNLI	Fake news detection	Text	Multi-Model	NELA-GT-19	~ 70% (AUC)
2024 [78]	Portable Graph Transformer	Rumor Detection	Text	Multi-Model	Snopes, Anomaly, Covid 19 news, Twitter 15/16, PHEME	76.9% (Acc) Snopes, 71.2% (Acc) Anomaly, 56.4% (Acc) Covid 19 news, 90.8% (Acc) Twitter 15/16, 75.9% (Acc) PHEME
2024 [79]	Bert, LSTM + VAE-Based encoders, IB principle	Sentiment Analysis	Text, Audio, Vision	Multi-Model	MOSI, MOSEI	87.2% (F1) 87% (Accuracy for MOSI) 85.8% (F1) 86.7% (Accuracy for MOSEI)
2024 [80]	CrossAMF, XLNet-CNN-BiGRU & reduced VGGNet	Automatic depression prediction	Text, Image, user behavior	Multi-Model	Zhejiang Provincial People's Hospital	89.15% (F1) 90.97% (Accuracy)

inconsistencies between audio and video features. Classifiers used were GMM (Gaussian Mixture Model), SVM, MLP, and LSTM. Krishnamurthy et al. [68] performed deception detection over a small dataset of 121 courtroom videos. They used Text-CNN for textual analysis, 3D-CNN for videos, and openSMILE with MLP for audio analysis. They also utilized micro-expression features such as smile, laughter, frown as another modality for deception detection. Data fusion techniques used were Concatenation and Hadamard + Concatenation. Wu et al. [87] and Rosas et al. [88] have also proposed deception detection on videos using real-life trial data.

Li et al. [89] detected tampered faces in videos to detect swapped faces of celebrities by detecting eye-blinking features. Eye blinking is detected using the LRCN (Long-term Recurrent Convolutional Networks) model, measuring how much open an eye is measured concerning frame coordinates. Features have been extracted using the VGG16 convolutional layer and propagated to a sequence learning module that uses LSTM-RNN, which can retain memory. Then, the probability is determined of how much the eye is open or closed. Bestagini et al. [90] detected local tampering in video sequences. This was done by finding duplicated frames in videos and cross-correlating them with Spatio-temporal frame regions.

4 Review methodology

The domain of fake news detection gained popularity very quickly. Large amounts of unverified and uncredible posts have been misleading people. Using linguistic features for credibility assessment of content is a popular and widely used method. Here, we list the existing challenges to detect fake news spreading through all types of data.

1. **Visual Fake News Detection:** The fake news menace is rising. It began with spreading through text and has now started gripping users through all forms of multimedia. Visual data probable to fake news exploitation can be categorized into images or videos. There are various existing approaches to detect text-based fake news. However, visuals play a great role in impacting viewers' minds and therefore are being infiltrated with fake news in the current generation. An image or video can be easily modified using media editing applications. Various manipulations in visual data go unrecognized through the viewers' eyes. It is very difficult for humans to observe minor changes in modified images and videos to classify them as fake or real. Automated tools are required which can identify minute variations created made to fabricate or manipulate visual data. This poses a great challenge for researchers in designing visual-based fake news detectors.
2. **Auditory Fake News Detection:** Auditory fake news is a type of fake news that has been in existence but is going unnoticed. Many social applications allow users to share recorded audios. These audios files are vulnerable to spreading fake content, propaganda, unverified information, and more. This type of multimedia has not been put into use for credibility assessment. There are no fake news detection mechanisms that incorporate audio as a sole modality or in combination with other data modalities. This issue needs to be addressed to prevent the contamination of audios with false content and serve its early detection.
3. **Detection of Embedded Fake Content:** When one type or modality of data is fixed or embedded within another type of data, it is embedded content. A new type of social media posts is spreading widely known as a 'meme.' It is an image or a video, mostly with text embedded on it. Various forms of media like text, image video, gifs, or hyperlinks are embedded into other forms. It is a complex task to analyze media that is embedded in some other data type. There is an upsurge of fake content in the form of embedded media or memes. Efficient detection mechanisms are required to fight such misleading content.
4. **Multimodal Datasets:** To build fake news detection mechanisms, most machine learning and deep learning tasks require large amounts of data. There is a lack of real-world multimodal datasets. Text-based fake news datasets are more in number than visual or multimodal datasets. Lack of proper datasets limits the extent of research. There is a need to collect real-world fake news data that consists of various types of information like text, image, video, and metadata.
5. **Holistic Detection Mechanism:** The research community has encountered many techniques that can robustly detect fake news. These techniques use linguistic features, visual features, sentiment scores, social context, network/propagation-based features, metadata, and hybrid features. There is no such mechanism at present that extracts all these details from a given fake content and predicts its integrity based on all the contributing factors. Different researchers have highlighted the importance of all of these techniques individually or in hybrid combinations. It is worthwhile to consider all these features for building a holistic fake news detector.
6. **Real-Time Verification:** Provided information-spread ease through the internet and online social platforms, fake news is being generated and spread at every instant. Fake news can be about anything and anyone. It spreads continuously as we interact on the internet. Existing detection tools either require users to self-validate a piece of news by fact-checking on their website/application or classify

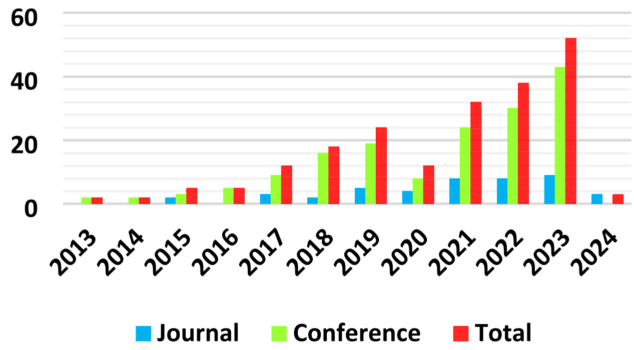

Fig. 4 Year-wise trend of published work

news late after it has been spread and affected various aspects of life. The world needs a system that analyses content in real-time and declares it as fake or real based on its decision.

7. **Lack of Literature:** There is a lack of significant literature in the domain of multi-modal fake news detection. Although many authors have presented textual detection mechanisms, work done in the multimodal domain is minimal and not complete. In this paper, we endeavor to cover all the past research performed using multiple data modalities other than only text-based techniques. We highlight works that have used visual content, alone or along with textual content, for detecting fake news.

After considering the challenges to detect fake news spreading through all types of data, we provide an enriched and systematic literature review of all the relevant and available articles in the multi-modal and visual fake news detection domain. The articles have been extracted from reputed digital libraries that include Google Scholar, IEEE Xplore, Elsevier, Springer, ACM, and Wiley.

We studied over 300 articles on the topic from the libraries mentioned above and shortlisted 202 relevant articles suitable for this survey that display state-of-the-art performance. We have covered all important articles and literature published in various journals, conferences, book chapters, and student thesis reports. Figure 4 provides an overview of the trend of works published yearly between 2013 and 2024. All the related research has been performed, and relevant articles have been published in this span. Some of the brilliant works done have been published in reputed journals or conferences. From Fig. 4, we can observe that focuses moved toward visual and multi-modal fake news detection largely after 2016. The number of published works has been rising since then. It is easily recognizable that tasks involving visual data in the fake news domain has been rising for the past five years and is grabbing researchers' attention.

The search words applied for querying digital databases include multi-modal fake news detection, image fake news detection, fake news detection, multi-modal fake news datasets, and their synonyms. Figure 5 presents the article-wise percentage distribution of algorithms and techniques utilized for multi-modal fake news detection.

5 Fake News Modalities

Fake news is defined as any piece of false information that misleads people. It can be deliberate, fabricated, or simply unintentional. The intent of spreading false news could be maliciously intentional, political, for gaining monetary benefits, popularity, or simply for fun. While referring to data modalities, fake news spreads through text, images, videos, audios, hyperlinks, embedded content, and hybrids. Because of less or no work in the remaining modalities, research is limited to textual and visual modalities. Therefore, we consider these modalities for review, which have been explored by researchers for fake news detection.

Text: This is the most popular mode of communication on the internet. People interact through textual matter on social media platforms, websites, blogs, e-mails, personal messaging, and more. Most of the false information spreads through text on the internet. Fake news is found propagating on social media posts, articles, and online messaging services. Text is the simplest and most used way for an internet user to convey his concerns. Being a largely used modality for communication, it also accounts for a large amount of fake news. Figure 6a shows an example of fake textual news. The screengrab is taken from Twitter. In the tweet, the user falsely attributes a claim to WION News, which states that China is hiding the real numbers of death amidst the coronavirus pandemic [91]. The post says that SO2 concentration around Wuhan, China, has grown due to the burning of a large number of dead bodies. The claim is false and has been debunked by various fact-checking websites.

Image: An image is a visual representation of something. Images are highly vivacious and impactful for depicting anything. They leave lasting impressions on the minds of viewers, whereas words can be forgotten shortly. They have suddenly gained popularity with the increase in the feasibility of sharing them. Images go through certain manipulations to carry a false message. The use of photo editing tools supports these manipulations. Some examples of editing techniques are cropping, splicing, copy-move, retouching, or blurring. Any image can be manipulated to convey a false message, which contributes to fake news. Often, they are not manipulated but accompanied by false text. Many of the times, irrelevant or out-of-context image is placed with fake text. All these types of images emanate false conceptions accord with fake news. In Fig. 6b, a Facebook post shows a girl rescuing a koala bear from Australia's bushfires [92]. Originally,

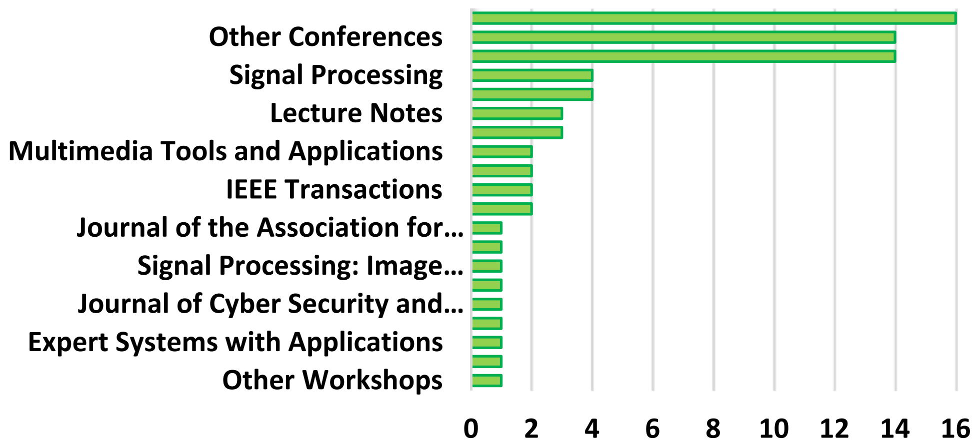

Fig. 5 Number of articles published in journals, conferences, and lecture notes

Fig. 6 **a** Textual fake news example [91] (Source: Twitter), **b** Example of fake news image [92] (Source: Facebook), **c** Fake example of a text-embedded image [93] (Source: Facebook)

the picture is a digitally created artwork used out-of-context to match the bushfire situation.

Video: Sharing of videos online became immensely easy with the introduction of YouTube. The platform allows the feasible sharing of information through video content. Social media platforms allow various techniques for communication through videos. This can be done in the form of regular posts, stories, ads, or even comments. This viability of video interaction gives room for sharing of fake content through videos. Videos are a powerful and impactful tool. They are capable of successfully manipulating people through their content. Therefore, it raises a serious concern to authenticate video content and decide whether a video is credible or not.

Other Modalities: Numerous data types have not been analyzed for fake news yet due to limitations of exposure and datasets. These include embedded data types, audios, and hyperlinks (clickbait). Embedded media is that where one type of data is merged or superimposed onto another. Textual matter on images or videos, embedding audio in images, altering audio in videos, etc. are some of examples. Detection of fraudulent content in such a data type is complex and challenging. There is a lack of past research in this area. Figure 6c shows a meme with a text embedded on it that says that the North Korean leader Kim Jong Un faked his death to expose traitors. Many such false statements and claims circled the internet [93]. Various fact-checking sites have debunked these claims.

6 Detection Mechanisms

Various fake news detection algorithms utilized by researchers are explained below. Figure 7 depicts the percentage distribution of used algorithms and methods in the reviewed articles. These include Deep Neural Networks, Image/Video Forensics, explicit features, and other methods. Textual detection mechanisms have employed several machine learning and deep learning classifiers. They have been successful enough in this domain. For visual analysis, the most popular tools are Deep Neural Networks. Many researchers have combined these with the use of explicit features available on the web.

These can be statistical features, user-based features, post-based features, propagation features, or more. Another

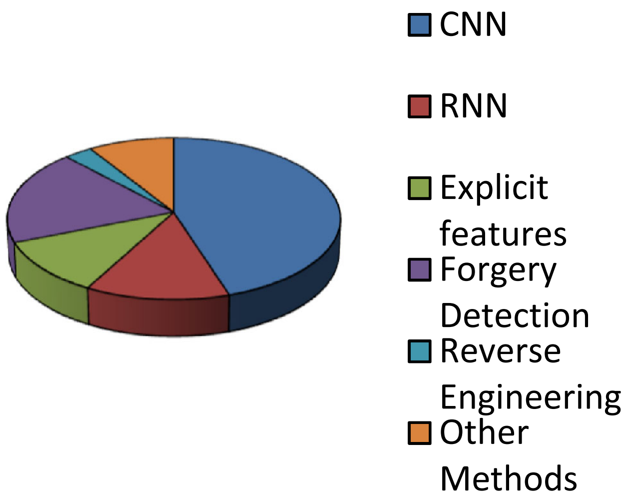

Fig. 7 Distribution of reviewed articles representing each method

popular method is the use of reverse image search on search engines. This method is useful in identifying the integrity of a particular image. Tools are available that use this mechanism to allow the verification of fake news. Several other methods of image and video manipulation have started being merged with fake news detection. This has bridged the gap between methodologies and brought the problem domain and possible solutions in a common range. The applications, advantages, and disadvantages of the following methods have been summarized in Table 3.

6.1 CNN

First introduced in the 1980s, CNNs have come a long way in the domain of computer vision. They have been applied to Natural Language Processing [94], image classification [95], video classification [96], object recognition [97], time-series forecasting [98], anomaly detection [99], speech analysis [100], handwriting recognition [101] and the likes. For image classification, CNNs require training over large image datasets. Their learning process occurred to be substantially faster than previous methods known, an underlying feature that brought CNNs into the picture. They are efficient in analyzing latent features present inside an image or a video. A rich survey of the latest Convolutional Neural Networks has been provided in Khan et al. [102]. As compared to images, little work has been done in video classification using CNNs as videos are more complex to process owing to their temporal dimension. Most works utilize CNNs to classify videos by extracting images at every video frame [68]. Another method treats spatial and temporal domain separately and classify them using two convolutional neural networks and fusing them after that. Many researchers have also applied CNNs for text classification using one-dimensional convolutional networks.

6.2 RNN

Recurrent Neural Networks analyze sequential inputs like text, image, speech, video, and output in a feedback loop. A network with feedback loops is created, which allows RNNs to retain information and train themselves. RNNs have been utilized in image classification [103], video classification [104], object recognition [105], video annotation [106], time-series prediction [107], anomaly detection [108], sentiment analysis [109], speech recognition and other ML and DL tasks. LSTMs (Long Short-Term Memory networks), a special case of RNNs, have been widely utilized in fake news classification tasks in multiple modalities. These train much faster and can perform complex classification tasks than other RNNs. 13% of the articles reviewed in this paper performing multi-modal fake news detection have utilized RNNs and a combination of RNNs and CNNs.

6.3 Reverse Engineering:

A popular method for fake news detection is to look up the content in question on the internet. Post's credibility is assured by matching it with the occurrences that appear in the search results. The methodology used under reverse engineering for fake news detection is Reverse Image Search. Search engines like Google, Yahoo, and Bing allow their users to input a query image and provide them with relevant information about the image. This process is utilized in fact-checking the authenticity of an image. We can get to know how old an image dates back and where did it appear first. Metadata can also be extracted from such visual data. It also helps us verify the context of the image to the text it accompanies. This method is used by automated fake news detection tools, applications, or web plugins.

Early attempts at catching fake news relied on the unsophisticated approaches: keyword matching or rule-based models, which were inefficient against more advanced disinformation campaigns. Table 3 represents Comparative analysis of fake news detection method (year in which these are introduced) represented methodologically. Traditional machine learning models used Naive Bayes and Support Vector Machines although giving a few percentage points of improvement were not able to distinguish between infinitesimal fine-tuning differences among subtle nuances in language and evolving strategies deployed by producers

Table 3 Comparative analysis of fake news detection methods

[Reference]	Method (Year)	Domain	Key Features	Applications	Advantages	Disadvantages
[94, 95]	Convolutional Neural Networks (**1980**)	Deep Learning	Convolutional layer- > Feature extraction, Pooling layer- > dimensionality reduction, FC layer- > classification & regression	Image, video, object, speech, time-series, handwritten data recognition, text mining	Adaptively learn classification features	Cannot detect image manipulation on its own Require large, labeled datasets
[103, 105, 106]	Recurrent Neural Networks (**1980s**)	Deep Learning	Sequential processing- > input data, Memory cell- > captures long term dependencies	Time, sequence-based predictions, Text, speech, image, video processing, object recognition	Great memorizing capacity, usable with CNNs	Slow computation, difficult training, hard to process long sequences, exploding and vanishing gradient
[110, 111]	Reverse Image Search (**2010**)	Reverse Engineering	Image hashing, feature extraction, similarity measures	Image origin tracking, metadata tracking, copyright infringement detection	Allows verifying variations in images by matching with originally occurring images and text	Can work with images previously present, cannot detect every type of fake image
[28, 112]	Statistical Analysis (**1980s**)	Explicit Feature Analysis	Quantitative analysis- > identify pattern & trends, Hypothesis testing, correction analysis, regression analysis	Explicit feature detection in all modalities based on features present out of the content, data exploration modeling	Provides prominent non-data features, measures semantics and relevance	Does not include latent features
[113, 114]	Propagation feature analysis (**1990s**)	Explicit Feature Analysis	Analyze the spread of information, Identify- > influential nodes & communities,	Explicit feature detection, rumor detection	Provides prominent non-data features	Assumption in normality
[49, 115]	Semantic score analysis (**1980–1990**)	Explicit Feature Analysis	Sentiment analysis, Topic modeling, text classification	Explicit feature detection, customer feedback analysis, Information retrieval	measures semantics	Does not include latent features
[38, 116]	User profile feature analysis (**2000s**)	Explicit Feature Analysis	Analyze user demographic, preference& behavior,	Fraud detection, targeted advertising, personalized recommendation	Provides prominent non-data features,	Subjectivity in factor rotation
[117, 118]	Geolocation, Psycholinguistics (**2000s**)	Explicit Feature Analysis	Studies how language use varies across different regions	Explicit feature detection, helps in understanding, cultural difference, regional sentiments, dialect variation	Provides prominent non-data features, and relevance	Does not include latent features
[119, 120]	Face manipulation detection (**2010–2015**)	Data Forensics	Detects alterations made to facial images	face-swap detection, deepfake detection, editing detection	Easily detects changes made in an image	Difficult to detect face manipulations in minutes
[121, 122]	Image/video tampering detection (**2000s**)	Data Forensics	Detect alterations made to images	Forged image and video detection, editing detection	Easily detects changes made in an image area by editing, removal, addition etc	Difficult to detect minute manipulations

of false content. These deep learning architectures, namely Recurrent Neural Networks and Convolutional Neural Networks, have shown extreme promise in discovering intricate patterns with both textual and visual data.

However, such an approach does consume significant amounts of high-quality data and is computationally expensive. Indeed, within the last couple of years, innovations such as transformer-based structures like BERT and its variants have basically transformed the domain. These models already proved to have high capability for handling contextual and semantic connotations and, thus, enable finding of subtle linguistic features that distinguish between news and information harming society and trustworthy information. Applicability of multimodal methodologies into the analysis of text, images, and videos opens frontiers for scholars to develop resilient systems designed with the aim of efficient detection and counteraction of misinfodissemination.

6.4 Explicit Features:

Under explicit features utilized for FND tasks, we categorize statistical features (no. of words, likes, shares, retweets, comments, reactions, etc.), similarity features that analyze the similarities between content and visual information of a news article and state how well both of these are correlated, semantic features that verify meaningfulness of data, user profile features that provide information about users' age groups, backgrounds, faiths and beliefs, inclination, online social behavior, and other relevant profile information, propagation features that help analyze the flow of fake news among networks and people, geolocation features those study areas of fake news generation and propagation and other external features. These features, when combined with other modalities, increase the weightage of detection accuracies. They serve as an important factor for fake news analysis and detection.

6.5 Data Forensics

Images and videos, given the current technological advancements, can be easily edited and tampered with. We have classified fake news detection techniques using forgery detection, splice-detection, copy-move detection, face-swapping detection, face manipulation detection, pixel-based forgery detection, Photoshop detection, object-removal detection, repurposing detection, and other similar editing detections under image forensics. This method verifies the credibility of images and videos without a requirement of their original version. Algorithms utilized can detect manipulated regions in images and videos. Popular literature that uses data forensics techniques for image and video classification is explained in Section 2.2.

6.6 Other Methods

Few other methods that have been utilized by researchers for fake news detection can be named as co-occurrence matrices, blockchain, pattern recognition, etc. It has become popular to match the semantics between post text, image, and video. Few of the latest works verify if the post's modalities convey the same meaning and then classify them as real or fake. They have provided a new dimension to investigate fake news detection. This area provides opportunities to be explored more, enhance currently available methods, and leverage new ones.

7 Trends across Modalities

This section presents an overview of crucial research performed in visual and multi-modal fake news detection. We highlight the vast usage of Deep Neural Networks and forgery detection techniques in multi-modal analysis through the survey. We present the survey classified based on the modalities used for fake news detection.

Qureshi and Deriche [123] explained the taxonomy of types of forgeries found in images: copy-move forgery, image retouching, resampling, image splicing. They have also discussed pixel-based forgery detection methods in images that include contrast enhancement detection, sharpening filtering detection, median filtering detection, resampling detection, post-processing editing detection, copy-move detection, and image splicing detection. Brezeale and Cook [124] provided a survey of existing video classification methods that classify videos using text features, audio features, visual features, and combinations. Boididou et al. [125] have reviewed various methodologies for classifying multimedia data on Twitter that include verifying cues, assessing the source and user credibility, content credibility, image forensics, verifying multimedia use task and have described the verification approaches used, namely UoS-ITI, MCG-ICT, and CERTH-UNITN. Anoop et al. [126] deeply studied fake news detection methods on textual modality, image modality, network modality, temporal modality, and knowledge-based approaches. They have also discussed popular datasets for use. Tolosana et al. [119] reviewed existing face-manipulation detection methodologies. Saini et al. [127] neatly summarized supervised, semi-supervised, and unsupervised multi-modal FND frameworks that include baselines like MVAE and EANN. They have compared state-of-the-art by nicely tabulated data.

7.1 Benchmark Datasets

Lack of suitable multi-modal datasets have, a lot, hampered the progress in the direction of fake news detection. Deep learning algorithms largely depend on huge amounts of training data, which, being meager, has appeared as a big challenge. The maximum number of fake news detection frameworks built till now have been trained upon data extracted from Twitter, Sina Weibo, or some websites.

For the time being, we present a piece of tabulated information about available datasets (image, video, and multi-modal) in chronological format. These have been used in the above-reviewed articles for fake news detection and similar tasks (Table 4). We also list out such datasets that contain news article URL or image/video URL [45]. These datasets can be further improved by extracting visual data using web scraping methods. A few of the other small-sized datasets have been generated for image analysis, which still is limited in number and not of optimum quality. Video datasets for this task are very rare and contain videos in the count of 100–200. Other video datasets are not completely relevant. There is an urgent demand for good quality multi-modal datasets that would furnish the need of the hour [46] [132]. The advancements in data augmentation or computer-generated data are beginning to contribute toward building datasets.

7.2 Performance Comparison

In this section, we demonstrate the usage of evaluation metrics utilized by fake news detection tasks and compare their performances based on the most utilized metrics, i.e., accuracy and F1-score. The comparison provided here is irrespective of the dataset but highlights each task's features and methods. We determine how the results have been moving all these years and identify prospective detection methods. Performances are displayed for tasks displaying the results achieved by the experiments on datasets they have used. Visual representations are provided for an easy understanding of how a given model performs when they use a specific set of features. Several metrics and parameters have been developed to define the functional performance or, in simpler terms, the algorithms' efficiency on giving the desired output of classification of data from a given dataset. Among them and widely utilized and relied upon metrics are Accuracy (in percentage), Precision, Recall, and F1 values, among several others such as AUC and EER. Almost all major research related to Fake News detection utilizes one or more among the former set of four metrics (Accuracy, Precision, Recall, and F-Score).

Hence, we bring forth such metric evaluation summary of the most relevant and pivotal experiments conducted for fake news detection. Figure 8 demonstrates the results of tasks in terms of F1-scores. We observe that the overall performance stays between 80–95% for methods that use textual and visual features combined.

The video classification task, which uses the annotation technique, still has a long way to go. In terms of accuracy, Fig. 9 we observe that the range of results is between 70–100%, with an average score of 85%. The majority of fake news classification tasks have relied upon deep neural networks. With the changing time, we also notice an inclination toward forensic algorithms for the same. Trends determine that most of the existing approaches have preferred to use deep learning algorithms due to their efficiency, robust nature, feasibility, and accuracy. Most works have preferred to use more than one feature, i.e., using multi-modal data. Thus, depending on more options and the type of fake information posts can offer. The aim is to consider all parameters that form/alters a user's perception of a piece of information. Convolutional Neural Networks, with maximum usage in the reviewed articles, have displayed eminent classification performance by exploiting the implicit features. They hold the potential to provide better results in future implementations.

8 Potential Research Directions

Research done in the past years is overwhelming yet, insufficient to cope with the amount of fake news pouring in. Each new happening or event in the world serves as a topic for fake news generation and propagation. In the present scenario, while a pandemic is going on, fake news reaches out to people more swiftly than authentic news is. No data modality is left behind in the race of spreading fake news. Text is not the only type of data of which one should be aware while intaking. People need to be more careful while digesting anything available on the internet because false information could greet us in any form, be it text, image, or video. So is the need for designing efficient and robust detection mechanisms. Analyzing the limitations and research gaps, this section highlights the potential directions where research can be proceeded into.

1. **Datasets:** From the above analysis, it is obvious that there is a shortage of large-scale multimodal datasets. Machine learning and Deep learning algorithms are data-driven. With the shortage of benchmark datasets, it becomes challenging to build detection mechanisms and compare various techniques' performances.

Table 4 Benchmark Multi-modal datasets

Dataset	Year	Type	Description	Class	Size	Source
MediaEval[a]	2015	Tweet, Image	Tweets related to 11 events (dev set) and 17 events (test set)	Binary	6,225 real tweets 9,596 fake tweets	Topsy, Twitter API
Mediaeval 2016[a] [30]	2016	Text, Image	Tweets and images from 53 past events	Binary	17,857 Tweets with 10,628 fake and 7,229 Real	Twitter API
FVC[a] [30]	2017	Text, Video	Fake Video Corpus for fake video detection	Binary	55 Fake and 49 Real videos	YouTube
Twitter [31]	2017	Text, Image	Text, Image, Social Context	Binary	6,026 Real and 7,898 Fake	Twitter
SinaWeibo [31]	2017	Text, Image	Data extracted from Chinese online social platform	Binary	4,749 Fake and 4,779 Real articles	Sina Weibo
Crawler News [37]	2017	Text, Image	News articles and images	Binary	2,500 Images	Google News
FakeNewsNet[a]	2018	Text, Image	News Content, Social Context, Dynamic Information, article URLs	Binary	6,000 Fake and 18,000 Real	PolitiFact, GossipCop
MFN [49]	2018	Text, Image	Event Centric dataset of tweets and corresponding images	Binary	14,000 Tweets and 500 Images 1,154 Real and 1,154 Fake News Articles	Twitter, Snopes and Web hose
Fakeddit[a] [128]	2019	Text, Image	Text, Image, Metadata and Comment	Multi-class	8 Lakh samples	Reddit
Politifact [23]	2019	Text, Image	News Content and their corresponding images, Retweet Comments	Binary	432 Fake and 624 Real News	Twitter
Gossip Cop [23]	2019	Text, Image	News Content and their corresponding images, Retweet Comments	Binary	5,323 fake 16,817 real	Magazines, Newspapers and social media
TI – CNN[a] [34]	2019	Text, Image	Text, metadata and Image URLs	Binary	20,015 total articles with 8,074 Real and 11,941 Fake	Over 240 Websites
[43]	2019	Text, Image	Metadata, News Articles	Binary	3,568 Fake, 15,915 Real	Multiple Websites
Newsbag [129]	2020	Text, Image	News articles and images	Binary	15,000 fake and 2 lakhs real	The Onion, The Wall Street Journal
News400 Dataset [53]	2020	Text, Image	Tweets, Articles	Binary	400	Twitter, Websites
Newsbag + + [129]	2020	Text, Image	Created by Data Augmentation	Binary	3,89,000 Fake and 2,00,000 Real	The Onion, The Wall Street Journal
ReCOVery Dataset [130]	2020	Text, Image, Social Information	News Articles, Tweets related to CovID-19	Binary	2,029 News Articles, 1,40,820 Tweets	Twitter, Multiple Websites

Table 4 (continued)

Dataset	Year	Type	Description	Class	Size	Source
Tampered News [53]	2020	Text, Image	News articles and images	Binary	72,561 News Articles	Breaking News Dataset
WhatsApp Dataset [131]	2020	Image	Fact-checked images extracted from WhatsApp groups		66,808	WhatsApp
Weibo21 Dataset [45]	2022	text	Chinese Fake text	Binary	4488 Chinese Fake news, 4640 real news	Chinese twitter
CIC Dataset [46]	2023	text	Fake text news	Multi-class	180,000 labeled tweets	Politifact news

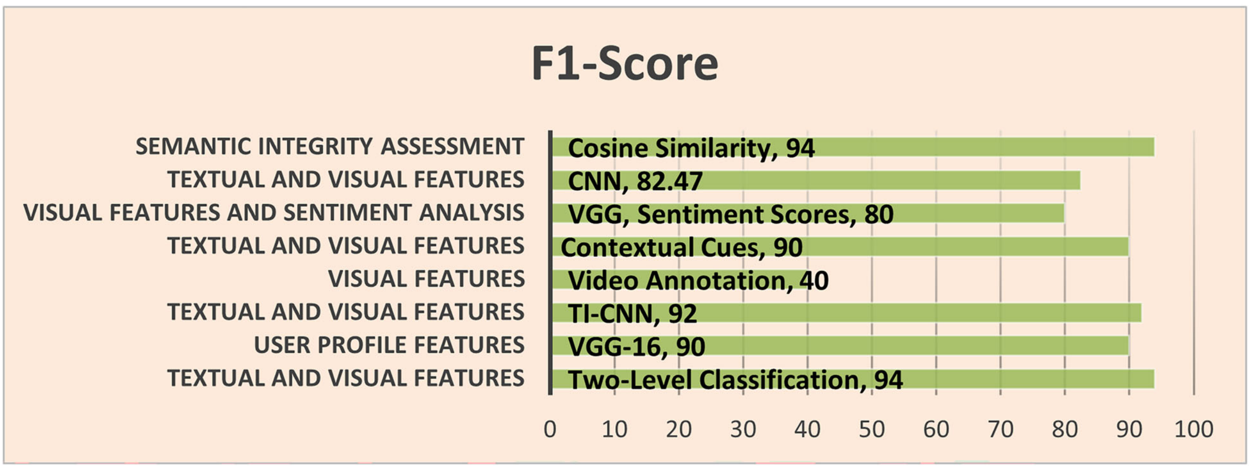

Fig. 8 Performance comparison based on F1-scores

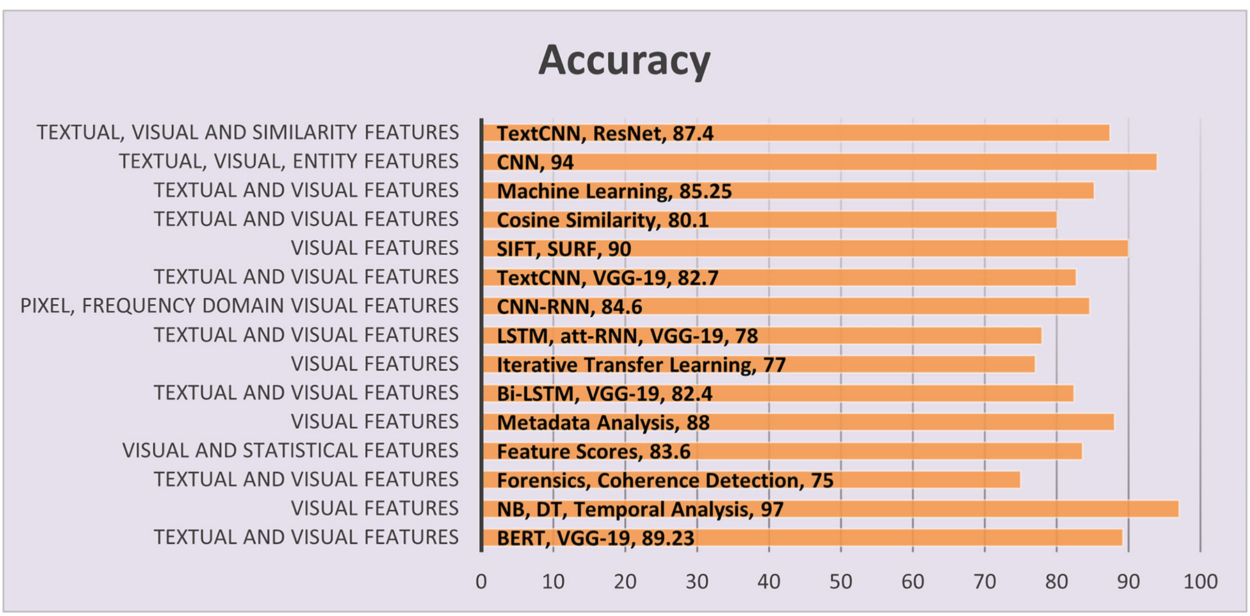

Fig. 9 Performance comparison based on accuracy scores

Although there are a few text datasets, those with multimodal information are limited and of poor quality. With the advantage of web-scraping mechanisms and free APIs, it has become easier to collect data. To proceed in the direction of multimodality, the collection of large-scale multimodal datasets is promoted.

2. **Real-time Detection:** With the assistance of deep learning algorithms, real-time detection models can be built to use fact-checked articles on the web for training and generate predictions for unseen data. There is a wide opportunity for the development of real-time detectors and automated fact-checkers.
3. **Early Detection:** Fake news detectors are built by feeding past data to algorithms. A baseline comparison is made on previous data. These algorithms are built when the fake information has already spread into the world and affected many. Entrapped within the fake news web, the world requires early detection of false news as and when it appears online. Users can only be benefited from fake news detectors when they provide early detection to prevent the propagation of fake news to a large scale. Early detection would allow intervention and thus mitigation of fake news before it spreads to a larger audience.
4. **Ubiquitous Detection Model:** With many social networking platforms available, it is challenging to incorporate a fake news detection mechanism to separate platforms individually. Similar content makes rounds on multiple platforms because one user can have accounts on various networks. This creates a replica of data on different social networks. With the help of redundant data and manual annotations, classification becomes easy for deep neural networks. A cross-platform system is required that would detect fake content on multiple social platforms. Implementation of models that can train on manually annotated content on one platform and then identify fake news on other platforms is suggested.
5. **Data-oriented Detection:** As far as the previous research is considered, we have very few frameworks that provide credibility assessment to fake content types. Most techniques consider text only, while some allow visual verification. It is challenging for a single system to verify the contents of all data modalities. Such a system would be more beneficial for the general public to authenticate information.
6. **Feature-oriented Detection:** All existing approaches use a limited subset of features, either linguistics, visuals, hybrid, data-centered, sentiment scores, social context, network-based, user-based, or post-based features. These contributing factors of fake news identification could be used all together for dependable predictions.
7. **Integrity Assessment:** In multimodal approaches, existing works perform detection based on features from each type of data independently. In many fake news instances, the post contents are not semantically related. The text, image, or video for a given post could be expressing unrelated context. Few works focus on assessing the semantic integrity of the news. This helps detect false news where data modalities have not been manipulated but are unrelated to each other. Such integrity assessment tools shall help in identifying out-of-context news items.
8. **Embedded Fake News Detection:** Detection of fake embedded content has not been done yet. A large volume of fake news is spreading through such type of data. To cope up with the incoming fake news, this type of detection mechanism is required.
9. **Multilingual Detection:** Current approaches have focused on English language data in text and videos. Due to the spread of fake news through regional languages on the web, multi-lingual approaches should be considered to detect fake news from other languages in the form of text, videos, or embedded content.
10. **Data Manipulation Detection:** With the popularity of image and video forensics techniques, forgery detection in data has become easier. Various manipulation techniques like face-spoofing detection, deepfake identification, tampering detection, splicing, copy-move detection, object removal/addition detection, etc., should be considered and merged with fake news detection mechanisms. There is a need for merging the domains of fake news detection and data manipulation detection.
11. **Browser Plugin/Application Software:** The availability of fake news detector tools in the form of easy-to-use browser plugins, add-ons, software, and mobile applications will enhance their accessibility and serve detection on a user-basis.

Concerning future works, we promote a multimodal framework that would efficiently detect fake news in all forms that revolve around the internet. We suggest exploring the domain incorporating fake news in the form of videos. We also motivate researchers to engage in building versatile multimodal datasets for future use collecting information from websites, online social platforms, and the likes. We encourage the readers to dive deeper into machine learning and deep learning algorithms and fish out ultimate solutions to the problem domain. Our work helps in bridging the research gaps and serve as potential future opportunities to work upon. We conclude this review anticipating that interested researchers will benefit from the information provided and narrow down their interests to this domain to contribute to the society and research community.

9 Conclusion

Uncontrolled and unauthentic data being over-loaded on the web needs appropriate solutions for the complexities being generated and has become a hard nut to crack. Deep learning algorithms are proving efficient and providing effective solutions with remarkable results. These solutions are to be unearthed from unimaginable horizons and that too within a very precise and limited period as the flow of complexities has reached the verge of parallel solutions. There has been a rapid increase in luring solutions for multimodal fake news detection adopting numerous variant techniques. This survey allows us to conclude that deep learning architectures prove astonishingly capable of fake news detection. They have resulted in high accuracies under the text-domain. Recurrent Neural Networks, LSTMs, GRU, Bidirectional GRU have contributed significantly to text classification. When it comes to visual data, Convolutional Neural Networks form the bigger picture. Survey displays that over 40% of methodologies have incorporated CNNs and their combinations with RNNs or other DNNs in their detection frameworks and served brilliant results. CNNs are taking the lead in computer vision, and allied domains and have become a prospective application for future FND tasks. Many researchers have identified fake images and videos and tampered regions in them, which we review as supportive tasks that can help classify fake news based on fake visuals. We motivate the readers to combine such tasks with FND modules to perform multimodal FND. By fusing modules performing such tasks on different modalities, optimized performances are assured.

There has been the unavailability of symbolic literature in this domain. The progress along the pathways of multimodal fake news detection has been slow. Researchers are unaware of the advancements so far reached. Existing literature is focused upon fake textual news and its detection mechanisms. This survey acknowledges this shortage and provides a broader and intact overview of multimodal fake news detection that incorporates image, video, audio, and their combinations with text. We segregate articles based on modalities implied in them. We demonstrate the year-wise trend of work involved and also analyze their method-wise distribution. We neatly summarize all the notable work and highlight important techniques that may pose powerful algorithms in future fake news classification tasks. Further, we described the evaluation metrics adapted in research so far.

We see that accuracy, precision, recall, and F-scores were observed for most of the tasks, whereas some liked to evaluate their models' performances using AUC and ROC. A few other metrics utilized were EER, HTER, TPR, and FPR. Accuracy appears as the most adopted method. Further, owing to the scarcity of multimodal datasets, we regard the obstacles faced in fledged research in the form of not so optimum solutions. Hence, we provide collective information of all the good quality image, video, and multimodal datasets available and previously been used in various tasks. This provides a route for fellow researchers to efficiently choose among the few available datasets and perform future research. We also encourage them to build good quality multimodal fake news datasets that would render this domain positively.

Author Contribution P.M. provided guidance and supervision throughout the research process. C.R. conducted the experiment. B. critically reviewed and revised the manuscript. Oversaw the submission and publication process.

Data Availability No datasets were generated or analyzed during the current study.

References

1. Zhang, Q., Guo, Z., Zhu, Y., Vijayakumar, P., Castiglione, A., Gupta, B.B.: A Deep Learning-based Fast Fake News Detection Model for Cyber-Physical Social Services. Pattern Recogn. Lett. **168**, 31–38 (2023). https://doi.org/10.1016/j.patrec.2023.02.026
2. Zhang, J., Tohidypour, H., Wang, Y., Nasiopoulos, P.: "Shallow- and Deep- fake Image Manipulation Localization Using Deep Learning," In: 2023 International Conference On Computing, Networking and Communications (Icnc), Honolulu, HI, USA: IEEE, (2023) https://doi.org/10.1109/ICNC57223.2023.10074246
3. Suratkar, S., Kazi, F.: Deep Fake Video Detection Using Transfer Learning Approach. Arab. J. Sci. Eng. **48**(8), 9727–9737 (2023). https://doi.org/10.1007/s13369-022-07321-3
4. Gramigna, R.: Preserving Anonymity: Deep-Fake as an Identity-Protection Device and as a Digital Camouflage. Int. J. Semiot. Law **37**(3), 729–751 (2024). https://doi.org/10.1007/s11196-023-10079-y
5. D. D'Cruze, "No! The government is doing no such thing': PIB team busts myth about govt reading WhatsApp messages," *BT Business Today*, 2023. [Online]. Available: https://www.businesstoday.in/technology/news/story/no-the-government-is-ng-no-such-thing-pib-team-busts-myth-about-govt-reading-whatsapp-messages-392317-2023-08-01
6. Archit and S. Satani, "Edited video of 'three-eyed' baby believed to be true on social media," *alt News*, 2020. [Online]. Available: https://www.altnews.in/edited-video-of-three-eyed-baby-believed-to-be-true-on-social-media/
7. "'Fact check: Viral picture of Facebook CEO Mark Zuckerberg, PM Narendra Modi is morphed.,'" *Newsmeter Network*, 2020. [Online]. Available: https://newsmeter.in/fact-check/fact-check-viral-picture-of-facebook-ceo-mark-zuckerberg-pm-narendra-modi-is-morphed-671882
8. "Kashmir fake photo: Fallout from the UN speech by Pakistan's Maleeha Lodhi," *BBC News*, 2017. [Online]. Available: https://www.bbc.com/news/world-asia-41427929
9. Yang, Z., Liang, J., Xu, Y., Zhang, X.-Y., He, R.: Masked Relation Learning for DeepFake Detection. IEEE Trans. Inform. Forensic Secur. **18**, 1696–1708 (2023). https://doi.org/10.1109/TIFS.2023.3249566
10. S. Sharma, "Joint multimodal sentiment analysis based on information relevance." Accessed: Dec. 12, 2023. [Online]. Available: https://ojs.aaai.org/index.php/ICWSM/article/view/7356/7210
11. Y. Ju, S. Hu, S. Jia, G. H. Chen, and S. Lyu, "Improving Fairness in Deepfake Detection".

12. Elkasrawi, S., Dengel, A., Abdelsamad, A., Bukhari, S.S.: “What You See is What You Get? Automatic Image Verification for Online News Content”, in 2016 12th IAPR Workshop on Document Analysis Systems (DAS). IEEE, Santorini, Greece (2016). https://doi.org/10.1109/DAS.2016.75
13. Y. Wang *et al.*, “EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection,” in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London United Kingdom: ACM, Jul. 2018, pp. 849–857. https://doi.org/10.1145/3219819.3219903.
14. S. Suryavardan *et al.*, “Factify 2: A Multimodal Fake News and Satire News Dataset”.
15. Ramin Safa, S.A., Edalatpanah, A.S.: Predicting mental health using social media: A roadmap for future development. In: Deep Learning in Personalized Healthcare and Decision Support, pp. 285–303. Elsevier (2023). https://doi.org/10.1016/B978-0-443-19413-9.00014-X
16. Junxia Ma, Lu., Rong, Y.Z., Tiwari, P.: Moving From Narrative to Interactive Multi-Modal Sentiment Analysis: A Survey. ACM Transactions on Asian and Low-Resource Language Information Processing (2023). https://doi.org/10.1145/3610288
17. “FakedBits- Detecting Fake Information on Social Platforms using Multi-Modal Features,” *KSII TIIS*, vol. 17, no. 1, Jan. 2023, https://doi.org/10.3837/tiis.2023.01.004.
18. Z. Jin, J. Cao, J. Luo, and Y. Zhang, “Image Credibility Analysis with Effective Domain Transferred Deep Networks,” Nov. 16, 2016, *arXiv*: arXiv:1611.05328. https://doi.org/10.48550/arXiv.1611.05328.
19. P. Qi, J. Cao, T. Yang, J. Guo, and J. Li, “Exploiting Multi-domain Visual Information for Fake News Detection,” in *2019 IEEE International Conference on Data Mining (ICDM)*, Beijing, China: IEEE, Nov. 2019, pp. 518–527. https://doi.org/10.1109/ICDM.2019.00062.
20. Vishwakarma, D.K., Varshney, D., Yadav, A.: Detection and veracity analysis of fake news via scrapping and authenticating the web search. Cogn. Syst. Res. **58**, 217–229 (2019). https://doi.org/10.1016/j.cogsys.2019.07.004
21. C. Pasquini, C. Brunetta, A. F. Vinci, V. Conotter, and G. Boato, “Towards the verification of image integrity in online news,” in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Turin, Italy: IEEE, Jun. 2015, pp. 1–6. https://doi.org/10.1109/ICMEW.2015.7169801.
22. Huh, M., Liu, A., Owens, A., Efros, A.A.: Fighting Fake News: Image Splice Detection via Learned Self-Consistency. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) Computer Vision – ECCV 2018: 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part XI, pp. 106–124. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-030-01252-6_7
23. L. Cui, S. Wang, and D. Lee, “SAME: sentiment-aware multi-modal embedding for detecting fake news,” in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Vancouver British Columbia Canada: ACM, Aug. 2019, pp. 41–48. https://doi.org/10.1145/3341161.3342894.
24. Z. Jin, J. Cao, Y. Zhang, and Y. Zhang, “MCG-ICT at MediaEval 2015: Verifying Multimedia Use with a Two-Level Classification Model”.
25. A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, “Faking Sandy: characterizing and identifying fake images on Twitter during Hurricane Sandy,” in *Proceedings of the 22nd International Conference on World Wide Web*, Rio de Janeiro Brazil: ACM, May 2013, pp. 729–736. https://doi.org/10.1145/2487788.2488033.
26. C. Maigrot, V. Claveau, E. Kijak, and R. Sicre, “MediaEval 2016: A multimodal system for the Verifying Multimedia Use task”.
27. Y. Zhang, L. Zheng, and V. L. L. Thing, “Automated face swapping and its detection,” in *2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP)*, Singapore: IEEE, Aug. 2017, pp. 15–19. https://doi.org/10.1109/SIPROCESS.2017.8124497.
28. Jin, Z., Cao, J., Zhang, Y., Zhou, J., Tian, Q.: Novel Visual and Statistical Image Features for Microblogs News Verification. IEEE Trans. Multimedia **19**(3), 598–608 (2017). https://doi.org/10.1109/TMM.2016.2617078
29. P. Zhou, X. Han, V. I. Morariu, and L. S. Davis, “Two-Stream Neural Networks for Tampered Face Detection,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA: IEEE, Jul. 2017, pp. 1831–1839. https://doi.org/10.1109/CVPRW.2017.229.
30. O. Papadopoulou, M. Zampoglou, S. Papadopoulos, and Y. Kompatsiaris, “Web Video Verification using Contextual Cues,” in *Proceedings of the 2nd International Workshop on Multimedia Forensics and Security*, Bucharest Romania: ACM, Jun. 2017, pp. 6–10. https://doi.org/10.1145/3078897.3080535.
31. Z. Jin, J. Cao, H. Guo, Y. Zhang, and J. Luo, “Multimodal Fusion with Recurrent Neural Networks for Rumor Detection on Microblogs,” in *Proceedings of the 25th ACM international conference on Multimedia*, Mountain View California USA: ACM, Oct. 2017, pp. 795–816. https://doi.org/10.1145/3123266.3123454.
32. Huckle, S., White, M.: Fake News: A Technological Approach to Proving the Origins of Content, Using Blockchains. Big Data **5**(4), 356–371 (2017). https://doi.org/10.1089/big.2017.0071
33. A. Jaiswal, E. Sabir, W. AbdAlmageed, and P. Natarajan, 2017 “Multimedia Semantic Integrity Assessment Using Joint Embedding Of Images And Text,” in *Proceedings of the 25th ACM international conference on Multimedia*, Mountain View California USA: ACM, https://doi.org/10.1145/3123266.3123385.
34. Y. Yang, L. Zheng, J. Zhang, Q. Cui, Z. Li, and P. S. Yu, “TI-CNN: Convolutional Neural Networks for Fake News Detection,” Jan. 13, 2023, *arXiv*: arXiv:1806.00749. https://doi.org/10.48550/arXiv.1806.00749.
35. S. Krishnan and M. Chen, “Identifying Tweets with Fake News,” in *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, Salt Lake City, UT: IEEE, Jul. 2018, pp. 460–464. https://doi.org/10.1109/IRI.2018.00073.
36. Yue, Wu., Abd-Almageed, W., Natarajan, P.: Busternet: Detecting copy-move image forgery with source/target localization. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) Computer Vision – ECCV 2018: 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part VI, pp. 170–186. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-030-01231-1_11
37. F. Lago, Q. Tin Phan, and G. Boato, “Visual and textual analysis for image trustworthiness assessment within online news,” *Security and Communication Networks*, no. 1, p. 9236910, 2019, https://doi.org/10.1155/2019/9236910.
38. K. Shu, X. Zhou, S. Wang, R. Zafarani, and H. Liu, “The role of user profiles for fake news detection,” in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Vancouver British Columbia Canada: ACM, Aug. 2019, pp. 436–439. https://doi.org/10.1145/3341161.3342927.
39. E. Sabir, J. Cheng, A. Jaiswal, W. AbdAlmageed, I. Masi, and P. Natarajan, “Recurrent Convolutional Strategies for Face Manipulation Detection in Videos,” in *Interfaces (GUI), 2019*, 2019, pp. 80–87. [Online]. Available: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://openaccess.thecvf.com/content_CVPRW_2019/papers/Media%20Forensics/Sabir_Recurrent_Convolutional_Strategies_for_

Face_Manipulation_Detection_in_Videos_CVPRW_2019_paper.pdf

40. D. Khattar, J. S. Goud, M. Gupta, and V. Varma, "MVAE: Multimodal Variational Autoencoder for Fake News Detection," in *The World Wide Web Conference*, San Francisco CA USA: ACM, May 2019, pp. 2915–2921. https://doi.org/10.1145/3308558.3313552.
41. Zhou, X., Jindi, Wu., Zafarani, R.: $$\mathsf {SAFE}$$: Similarity-Aware Multi-modal Fake News Detection. In: Lauw, H.W., Wong, R.-W., Ntoulas, A., Lim, E.-P., Ng, S.-K., Pan, S.J. (eds.) Advances in Knowledge Discovery and Data Mining: 24th Pacific-Asia Conference, PAKDD 2020, Singapore, May 11–14, 2020, Proceedings, Part II, pp. 354–367. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-47436-2_27
42. D. Zlatkova, P. Nakov, and I. Koychev, "Fact-Checking Meets Fauxtography: Verifying Claims About Images," Aug. 30, 2019, *arXiv*: arXiv:1908.11722. https://doi.org/10.48550/arXiv.1908.11722.
43. Singh, V.K., Ghosh, I., Sonagara, D.: Detecting fake news stories via multimodal analysis. Asso for Info Science & Tech **72**(1), 3–17 (2021). https://doi.org/10.1002/asi.24359
44. M. Steinebach, K. Gotkowski, and H. Liu, "Fake News Detection by Image Montage Recognition," in *Proceedings of the 14th International Conference on Availability, Reliability and Security*, Canterbury CA United Kingdom: ACM, Aug. 2019, pp. 1–9. https://doi.org/10.1145/3339252.3341487.
45. Jing, J., Wu, H., Sun, J., Fang, X., Zhang, H.: Multimodal fake news detection via progressive fusion networks. Inf. Process. Manage. **60**(1), 103120 (2023). https://doi.org/10.1016/j.ipm.2022.103120
46. Zhang, X., Dadkhah, S., Weismann, A.G., Kanaani, M.A., Ghorbani, A.A.: Multimodal Fake News Analysis Based on Image-Text Similarity. IEEE Trans. Comput. Soc. Syst. **11**(1), 959–972 (2024). https://doi.org/10.1109/TCSS.2023.3244068
47. Umair, M., Bouguettaya, A., Lakhdari, A., Ouzzani, M., Liu, Y.: Exif2Vec: A Framework to Ascertain Untrustworthy Crowd-sourced Images Using Metadata. ACM Trans. Web **18**(3), 1–27 (2024). https://doi.org/10.1145/3645094
48. O. Ajao, D. Bhowmik, and S. Zargari, "Fake News Identification on Twitter with Hybrid CNN and RNN Models," in *Proceedings of the 9th International Conference on Social Media and Society*, Copenhagen Denmark: ACM, Jul. 2018, pp. 226–230. https://doi.org/10.1145/3217804.3217917.
49. S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, and S. Satoh, "SpotFake: A Multi-modal Framework for Fake News Detection," in *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)*, Singapore, Singapore: IEEE, Sep. 2019, pp. 39–47. https://doi.org/10.1109/BigMM.2019.00-44.
50. D. Sáez-Trumper, "Fake tweet buster: a webtool to identify users promoting fake news on twitter," in *In Proceedings of the 25th ACM conference on Hypertext and social media*, Association for Computing Machinery, 2014, pp. 316–317. https://doi.org/10.1145/2631775.2631786.
51. G. Armano *et al.*, "NewsVallum: Semantics-Aware Text and Image Processing for Fake News Detection system⋆".
52. Chen, Y., Conroy, N.K., Rubin, V.L.: News in an online world: The need for an 'automatic crap detector.' Proc. Assoc. Info. Sci. Tech. **52**(1), 1–4 (2015). https://doi.org/10.1002/pra2.2015.145052010081
53. E. Müller-Budack, J. Theiner, S. Diering, M. Idahl, and R. Ewerth, "Multi-modal analytics for real-world news using measures of cross-modal entity consistency," in *Proceedings of the 2020 international conference on multimedia retrieval*, Dublin, Ireland: Association for Computing Machinery, pp. 16–25. https://doi.org/10.1145/3372278.3390670.
54. S. B. Parikh, S. R. Khedia, and P. K. Atrey, "A Framework to Detect Fake Tweet Images on Social Media," in *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)*, Singapore, Singapore: IEEE, Sep. 2019, pp. 104–110. https://doi.org/10.1109/BigMM.2019.00-37.
55. Nixon, L., Apostolidis, E., Markatopoulou, F., Patras, I., Mezaris, V.: Multimodal Video Annotation for Retrieval and Discovery of Newsworthy Video in a News Verification Scenario. In: Kompatsiaris, I., Huet, B., Mezaris, V., Gurrin, C., Cheng, W.-H., Vrochidis, S. (eds.) MultiMedia Modeling: 25th International Conference, MMM 2019, Thessaloniki, Greece, January 8–11, 2019, Proceedings, Part I, pp. 143–155. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-05710-7_12
56. A. Bagade *et al.*, "The Kauwa-Kaate Fake News Detection System: Demo," in *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, Hyderabad India: ACM, Jan. 2020, pp. 302–306. https://doi.org/10.1145/3371158.3371402.
57. Boididou, C., Papadopoulos, S., Zampoglou, M., Apostolidis, L., Papadopoulou, O., Kompatsiaris, Y.: Detection and visualization of misleading content on Twitter. Int J Multimed Info Retr **7**(1), 71–86 (2018). https://doi.org/10.1007/s13735-017-0143-x
58. S. Sun, H. Liu, J. He, and X. Du, "Detecting Event Rumors on Sina Weibo Automatically," in *Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer, pp. 120–131. https://doi.org/10.1007/978-3-642-37401-2_14.
59. B. Bayar and M. C. Stamm, "A Deep Learning Approach to Universal Image Manipulation Detection Using a New Convolutional Layer," in *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security*, Vigo Galicia Spain: ACM, Jun. 2016, pp. 5–10. https://doi.org/10.1145/2909827.2930786.
60. E. Sabir, W. AbdAlmageed, Y. Wu, and P. Natarajan, "Deep multimodal image-repurposing detection," in *MM '18: Proceedings of the 26th ACM international conference on Multimedia*, Seoul, Republic of Korea: Association for Computing Machinery, 2018, pp. 1337–1345. https://doi.org/10.1145/3240508.3240707.
61. A. Jaiswal, Y. Wu, W. AbdAlmageed, I. Masi, and P. Natarajan, "Aird: Adversarial learning framework for image repurposing detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, California, 2019, pp. 11330–11339.
62. T. Pomari, G. Ruppert, E. Rezende, A. Rocha, and T. Carvalho, "Image Splicing Detection Through Illumination Inconsistencies and Deep Learning," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens: IEEE, Oct. 2018, pp. 3788–3792. https://doi.org/10.1109/ICIP.2018.8451227.
63. M. Zampoglou, S. Papadopoulos, and Y. Kompatsiaris, "Detecting image splicing in the wild (WEB)," in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Turin, Italy: IEEE, Jun. 2015, pp. 1–6. https://doi.org/10.1109/ICMEW.2015.7169839.
64. Y. Wu, W. Abd-Almageed, and P. Natarajan, "Deep Matching and Validation Network: An End-to-End Solution to Constrained Image Splicing Localization and Detection," in *Proceedings of the 25th ACM international conference on Multimedia*, Mountain View California USA: ACM, Oct. 2017, pp. 1480–1502. https://doi.org/10.1145/3123266.3123411.
65. J. Yang, Z. Lei, and S. Z. Li, "Learn Convolutional Neural Network for Face Anti-Spoofing," Aug. 26, 2014, *arXiv*: arXiv:1408.5601. https://doi.org/10.48550/arXiv.1408.5601.
66. P. Korshunov and S. Marcel, "Speaker Inconsistency Detection in Tampered Video," in *2018 26th European Signal Processing Conference (EUSIPCO)*, Rome: IEEE, Sep. 2018, pp. 2375–2379. https://doi.org/10.23919/EUSIPCO.2018.8553270.
67. F. Marra, D. Gragnaniello, D. Cozzolino, and L. Verdoliva, "Detection of GAN-Generated Fake Images over Social Networks,"

in *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, Miami, FL: IEEE, Apr. 2018, pp. 384–389. https://doi.org/10.1109/MIPR.2018.00084.
68. Krishnamurthy, G., Majumder, N., Poria, S., Cambria, E.: A Deep Learning Approach for Multimodal Deception Detection. In: Gelbukh, A. (ed.) Computational Linguistics and Intelligent Text Processing: 19th International Conference, CICLing 2018, Hanoi, Vietnam, March 18–24, 2018, Revised Selected Papers, Part I, pp. 87–96. Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-23793-5_8
69. S. Tariq, S. Lee, H. Kim, Y. Shin, and S. S. Woo, "Detecting Both Machine and Human Created Fake Face Images In the Wild," in *Proceedings of the 2nd International Workshop on Multimedia Privacy and Security*, Toronto Canada: ACM, Jan. 2018, pp. 81–87. https://doi.org/10.1145/3267357.3267367.
70. H. H. Nguyen, T. N.-D. Tieu, H.-Q. Nguyen-Son, V. Nozick, J. Yamagishi, and I. Echizen, "Modular Convolutional Neural Network for Discriminating between Computer-Generated Images and Photographic Images," in *Proceedings of the 13th International Conference on Availability, Reliability and Security*, Hamburg Germany: ACM, Aug. 2018, pp. 1–10. https://doi.org/10.1145/3230833.3230863.
71. V. V. Kniaz, V. Knyaz, and F. Remondino, "The Point Where Reality Meets Fantasy: Mixed Adversarial Generators for Image Splice Detection".
72. D. Guera and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," in *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Auckland, New Zealand: IEEE, Nov. 2018, pp. 1–6. https://doi.org/10.1109/AVSS.2018.8639163.
73. H. H. Nguyen, J. Yamagishi, and I. Echizen, "Capsule-forensics: Using Capsule Networks to Detect Forged Images and Videos," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom: IEEE, May 2019, pp. 2307–2311. https://doi.org/10.1109/ICASSP.2019.8682602.
74. L. Nataraj *et al.*, "Detecting GAN generated Fake Images using Co-occurrence Matrices," Oct. 03, 2019, *arXiv*: arXiv:1903.06836. https://doi.org/10.48550/arXiv.1903.06836.
75. Li, H., Li, B., Tan, S., Huang, J.: Identification of deep network generated images using disparities in color components. Signal Process. **174**, 107616 (2020). https://doi.org/10.1016/j.sigpro.2020.107616
76. Xue, J., Wang, Y., Tian, Y., Li, Y., Shi, L., Wei, L.: Detecting fake news by exploring the consistency of multimodal data. Inf. Process. Manage. **58**(5), 102610 (2021). https://doi.org/10.1016/j.ipm.2021.102610
77. Raza, S., Ding, C.: Fake news detection based on news content and social contexts: a transformer-based approach. Int J Data Sci Anal **13**(4), 335–362 (2022). https://doi.org/10.1007/s41060-021-00302-z
78. Nguyen, T.T., Ren, Z., Nguyen, T.T., Jo, J., Nguyen, Q.V.H., Yin, H.: Portable graph-based rumour detection against multi-modal heterophily. Knowl.-Based Syst. **284**, 111310 (2024). https://doi.org/10.1016/j.knosys.2023.111310
79. Liu, W., Cao, S., Zhang, S.: Multimodal consistency-specificity fusion based on information bottleneck for sentiment analysis. Journal of King Saud University - Computer and Information Sciences **36**(2), 101943 (2024). https://doi.org/10.1016/j.jksuci.2024.101943
80. Wang, L., Zhang, Y., Zhou, B., Cao, S., Hu, K., Tan, Y.: Automatic depression prediction via cross-modal attention-based multi-modal fusion in social networks. Comput. Electr. Eng. **118**, 109413 (2024). https://doi.org/10.1016/j.compeleceng.2024.109413
81. N. Rahmouni, V. Nozick, J. Yamagishi, and I. Echizen, "Distinguishing computer graphics from natural images using convolution neural networks," in *2017 IEEE Workshop on Information Forensics and Security (WIFS)*, Rennes: IEEE, Dec. 2017, pp. 1–6. https://doi.org/10.1109/WIFS.2017.8267647.
82. De Rezende, E.R.S., Ruppert, G.C.S., Theóphilo, A., Tokuda, E.K., Carvalho, T.: Exposing computer generated images by using deep convolutional neural networks. Signal Processing: Image Communication **66**, 113–126 (2018). https://doi.org/10.1016/j.image.2018.04.006
83. Dang, L.M., Hassan, S.I., Im, S., Moon, H.: Face image manipulation detection based on a convolutional neural network. Expert Syst. Appl. **129**, 156–168 (2019). https://doi.org/10.1016/j.eswa.2019.04.005
84. Y. Wu, W. AbdAlmageed, and P. Natarajan, "ManTra-Net: Manipulation Tracing Network for Detection and Localization of Image Forgeries With Anomalous Features," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, CVF, 2019, pp. 9543–9552.
85. S.-Y. Wang, O. Wang, R. Zhang, A. Owens, and A. Efros, "Detecting Photoshopped Faces by Scripting Photoshop," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South): IEEE, Oct. 2019, pp. 10071–10080. https://doi.org/10.1109/ICCV.2019.01017.
86. P. Korshunov and S. Marcel, "DeepFakes: a New Threat to Face Recognition? Assessment and Detection," Dec. 20, 2018, *arXiv*: arXiv:1812.08685. https://doi.org/10.48550/arXiv.1812.08685.
87. Zhe, Wu., Singh, B., Larry Davis, V., Subrahmanian,: Deception Detection in Videos. Proceedings of the AAAI Conference on Artificial Intelligence (2018). https://doi.org/10.1609/aaai.v32i1.11502
88. V. Pérez-Rosas, M. Abouelenien, R. Mihalcea, and M. Burzo, "Deception Detection using Real-life Trial Data," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, Seattle Washington USA: ACM, Nov. 2015, pp. 59–66. https://doi.org/10.1145/2818346.2820758.
89. Y. Li, M.-C. Chang, and S. Lyu, "In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking," in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, Hong Kong, Hong Kong: IEEE, Dec. 2018, pp. 1–7. https://doi.org/10.1109/WIFS.2018.8630787.
90. P. Bestagini, S. Milani, M. Tagliasacchi, and S. Tubaro, "Local tampering detection in video sequences," in *2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSP)*, Pula (CA), Italy: IEEE, Sep. 2013, pp. 488–493. https://doi.org/10.1109/MMSP.2013.6659337.
91. *Gravitas: Is China hiding the real numbers of the Coronavirus outbreak?* [Online Video]. Available: https://www.youtube.com/watch?v=XaXeA3z2K7Q
92. B. Singh, "Fact Check: Did this girl rescue Koala bear from Australia Bushfire? No, this is a digitally created artwork," *Vishvas News*, 2020. [Online]. Available: https://www.vishvasnews.com/english/world/fact-check-did-this-girl-rescue-koala-bear-from-australia-bushfire-no-this-is-a-digitally-created-artwork/
93. S. Necessary and KETK, "Media reports say North Korean dictator Kim Jong Un dead," *Kron4*, San Fransisco, 2020. [Online]. Available: https://www.kron4.com/news/media-reports-say-north-korean-dictator-kim-jong-un-dead/
94. Widiastuti, N.I.: Convolution Neural Network for Text Mining and Natural Language Processing. IOP Conference Series: Materials Science and Engineering **662**(5), 052010 (2019). https://doi.org/10.1088/1757-899X/662/5/052010
95. A. A. M. Al-Saffar, H. Tao, and M. A. Talab, "Review of deep convolution neural network in image classification," in *2017 International Conference on Radar, Antenna, Microwave, Electronics,*

and Telecommunications (ICRAMET), Jakarta: IEEE, Oct. 2017, pp. 26–31. https://doi.org/10.1109/ICRAMET.2017.8253139.
96. A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-Scale Video Classification with Convolutional Neural Networks," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA: IEEE, Jun. 2014, pp. 1725–1732. https://doi.org/10.1109/CVPR.2014.223.
97. Singh, R.D., Mittal, A., Bhatia, R.K.: 3D convolutional neural network for object recognition: a review. Multimed Tools Appl **78**(12), 15951–15995 (2019). https://doi.org/10.1007/s11042-018-6912-6
98. Zhao, B., et al.: Convolutional neural networks for time series classification. JSEE **28**(1), 162–169 (2017). https://doi.org/10.21629/JSEE.2017.01.18
99. D. Kwon, K. Natarajan, S. C. Suh, H. Kim, and J. Kim, "An Empirical Study on Network Anomaly Detection Using Convolutional Neural Networks," in *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*, Vienna: IEEE, Jul. 2018, pp. 1595–1598. https://doi.org/10.1109/ICDCS.2018.00178.
100. J.-T. Huang, J. Li, and Y. Gong, "An analysis of convolutional neural networks for speech recognition," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, Queensland, Australia: IEEE, Apr. 2015, pp. 4989–4993. https://doi.org/10.1109/ICASSP.2015.7178920.
101. M. M. Abu Ghosh and A. Y. Maghari, "A Comparative Study on Handwriting Digit Recognition Using Neural Networks," in *2017 International Conference on Promising Electronic Technologies (ICPET)*, Deir El-Balah: IEEE, Oct. 2017, pp. 77–81. https://doi.org/10.1109/ICPET.2017.20.
102. Khan, A., Sohail, A., Zahoora, U., Qureshi, A.S.: A survey of the recent architectures of deep convolutional neural networks. Artif. Intell. Rev. **53**(8), 5455–5516 (2020). https://doi.org/10.1007/s10462-020-09825-6
103. Mou, L., Ghamisi, P., Zhu, X.X.: Deep Recurrent Neural Networks for Hyperspectral Image Classification. IEEE Trans. Geosci. Remote Sensing **55**(7), 3639–3655 (2017). https://doi.org/10.1109/TGRS.2016.2636241
104. Joe Yue-Hei Ng, M. Hausknecht, S. Vijayanarasimhan, O. Vinyals, R. Monga, and G. Toderici, "Beyond short snippets: Deep networks for video classification," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA: IEEE, Jun. 2015, pp. 4694–4702. https://doi.org/10.1109/CVPR.2015.7299101.
105. Ming Liang and Xiaolin Hu, "Recurrent convolutional neural network for object recognition," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA: IEEE, Jun. 2015, pp. 3367–3375. https://doi.org/10.1109/CVPR.2015.7298958.
106. P. Pan, Z. Xu, Y. Yang, F. Wu, and Y. Zhuang, "Hierarchical Recurrent Neural Encoder for Video Representation with Application to Captioning," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 1029–1038. https://doi.org/10.1109/CVPR.2016.117.
107. P. G. Madhavan, "Recurrent neural network for time series prediction," in *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Societ*, San Diego, CA: IEEE, 1993, pp. 250–251. https://doi.org/10.1109/IEMBS.1993.978527.
108. P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, "Long Short Term Memory Networks for Anomaly Detection in Time Series," *Computational Intelligence*, 2015.
109. J. Ruales, "Recurrent Neural Networks for Sentiment Analysis".
110. M. Moayeri *et al.*, "Rethinking Artistic Copyright Infringements in the Era of Text-to-Image Generative Models," Apr. 11, 2024, *arXiv*: arXiv:2404.08030. https://doi.org/10.48550/arXiv.2404.08030.
111. Aprin, F., Irene-Angelica Chounta, H., Hoppe, U.: "See the Image in Different Contexts": Using Reverse Image Search to Support the Identification of Fake News in Instagram-Like Social Media. In: Crossley, S., Popescu, E. (eds.) Intelligent Tutoring Systems: 18th International Conference, ITS 2022, Bucharest, Romania, June 29 – July 1, 2022, Proceedings, pp. 264–275. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-031-09680-8_25
112. Lima, G.B., Chaves, T.D.M., Freitas, W.W.L., De Souza, R.M.C.R.: Statistical learning from Brazilian fake news. Expert. Syst. **40**(3), e13171 (2023). https://doi.org/10.1111/exsy.13171
113. J. Zhu, C. Gao, Z. Yin, X. Li, and J. Kurths, "Propagation Structure-Aware Graph Transformer for Robust and Interpretable Fake News Detection," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Barcelona Spain: ACM, Aug. 2024, pp. 4652–4663. https://doi.org/10.1145/3637528.3672024.
114. Shu, K., Mahudeswaran, D., Wang, S., Liu, H.: Hierarchical Propagation Networks for Fake News Detection: Investigation and Exploitation. ICWSM **14**, 626–637 (2020). https://doi.org/10.1609/icwsm.v14i1.7329
115. Li, H., Cai, Z., Graesser, A.C.: Computerized summary scoring: crowdsourcing-based latent semantic analysis. Behav Res **50**(5), 2144–2161 (2018). https://doi.org/10.3758/s13428-017-0982-7
116. G. Antoniol and M. Di Penta, 2004 "A distributed architecture for dynamic analyses on user-profile data," in *Eighth European Conference on Software Maintenance and Reengineering, 2004. CSMR 2004. Proceedings.*, Tampere, Finland: IEEE, , pp. 319–328. https://doi.org/10.1109/CSMR.2004.1281434.
117. Tahmasebzadeh, G., Kacupaj, E., Müller-Budack, E., Hakimov, S., Lehmann, J., Ewerth, R.: GeoWINE: geolocation based wiki, image, news and event retrieval. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event Canada: ACM, Jul. 2021, pp. 2565–2569. https://doi.org/10.1145/3404835.3462786
118. Giachanou, A., Ghanem, B., Ríssola, E.A., Rosso, P., Crestani, F., Oberski, D.: The impact of psycholinguistic patterns in discriminating between fake news spreaders and fact checkers. Data Knowl. Eng. **138**, 101960 (2022). https://doi.org/10.1016/j.datak.2021.101960
119. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., Ortega-Garcia, J.: Deepfakes and beyond: A Survey of face manipulation and fake detection. Information Fusion **64**, 131–148 (2020). https://doi.org/10.1016/j.inffus.2020.06.014
120. Neves, J.C., Tolosana, R., Vera-Rodriguez, R., Lopes, V., Proenca, H., Fierrez, J.: GANprintR: Improved Fakes and Evaluation of the State of the Art in Face Manipulation Detection. IEEE J. Sel. Top. Signal Process. **14**(5), 1038–1048 (2020). https://doi.org/10.1109/JSTSP.2020.3007250
121. Akhtar, N., Saddique, M., Asghar, K., Bajwa, U.I., Hussain, M., Habib, Z.: Digital Video Tampering Detection and Localization: Review, Representations, Challenges and Algorithm. Mathematics **10**(2), 168 (2022). https://doi.org/10.3390/math10020168
122. Rocha, A., Scheirer, W., Boult, T., Goldenstein, S.: Vision of the unseen: Current trends and challenges in digital image and video forensics. ACM Comput. Surv. **43**(4), 1–42 (2011). https://doi.org/10.1145/1978802.1978805
123. Qureshi, M.A., Deriche, M.: A bibliography of pixel-based blind image forgery detection techniques. Signal Processing: Image Communication **39**, 46–74 (2015). https://doi.org/10.1016/j.image.2015.08.008
124. D. Brezeale and D. J. Cook, "Automatic Video Classification: A Survey of the Literature," *IEEE Trans. Syst., Man, Cybern. C*,

vol. 38, no. 3, pp. 416–430, May 2008, https://doi.org/10.1109/TSMCC.2008.919173.
125. Boididou, C., et al.: Verifying information with multimedia content on twitter: A comparative study of automated approaches. Multimed Tools Appl **77**(12), 15545–15571 (2018). https://doi.org/10.1007/s11042-017-5132-9
126. K. Anoop, M. P. Gangan, D. P, and V. L. Lajish, "Leveraging Heterogeneous Data for Fake News Detection," in *Linking and Mining Heterogeneous and Multi-view Data*, Springer, 2018. [Online]. Available: https://doi.org/10.1007/978-3-030-01872-6_10
127. N. Saini, M. Singhal, M. Tanwar, and P. Meel, "Multimodal, Semi-supervised and Unsupervised web content credibility analysis Frameworks," in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India: IEEE, May 2020, pp. 948–955. https://doi.org/10.1109/ICICCS48265.2020.9121005.
128. K. Nakamura, S. Levy, and W. Y. Wang, "r/Fakeddit: A New Multimodal Benchmark Dataset for Fine-grained Fake News Detection," Mar. 12, 2020, *arXiv*: arXiv:1911.03854. https://doi.org/10.48550/arXiv.1911.03854.
129. S. Jindal, R. Sood, R. Singh, M. Vatsa, and T. Chakraborty, "NewsBag: A Multimodal Benchmark Dataset for Fake News Detection".
130. X. Zhou, A. Mulay, E. Ferrara, and R. Zafarani, "ReCOVery: A Multimodal Repository for COVID-19 News Credibility Research," in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, Virtual Event Ireland: ACM, Oct. 2020, pp. 3205–3212. https://doi.org/10.1145/3340531.3412880.
131. J. C. S. Reis, P. Melo, K. Garimella, J. M. Almeida, D. Eckles, and F. Benevenuto, 2020 "A Dataset of Fact-Checked Images Shared on WhatsApp During the Brazilian and Indian Elections," in *Proceedings of the International AAAI Conference on Web and Social Media*, Archives, https://doi.org/10.1609/icwsm.v14i1.7356.
132. Chen, D., Su, W., Wu, P., Hua, B.: Joint multimodal sentiment analysis based on information relevance. Inf. Process. Manage. **60**(2), 103193 (2023). https://doi.org/10.1016/j.ipm.2022.103193

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