IJoSI

DOI: 10.6977/IJoSI.202306_7(6).0004
D. Nikumbh, A. Thakare/Int. J. Systematic Innovation, 7(6), 36-53(2023)

A comprehensive review of fake news detection on social media: feature engineering, feature fusion, and future research directions

Deepti Nikumbh[1*], Anuradha Thakare[2]
[1]Dept of Computer Engineering, Savitribai Phule Pune University, Pune, India.
[*] Corresponding author E-mail: Deepti.nikumbh@gmail.com
(Received 10 January 2023; Final version received 29 April 2023; Accepted 4 June 2023)

Abstract

Social media platforms are mainly used for information sharing, connecting with people, and staying updated about the latest events. However, information present on social media is sometimes incorrect, unverified, or misleading. Such information is often termed fake news. It is deliberately written to deceive the readers. It has the potential to change their perception of the topic or content being discussed. The best medium to share fake news is social media platforms. Large amounts of misleading or fake online information can have serious consequences. It can affect the social, political, economic well-being of individuals, society, and a nation as a whole. Fake News in the form of satire, fabricated and manipulated content, misleading information, and conspiracy theories get more likes and shares on social media and they spread quickly in no time. Thus, fake news detection (FND) and prevention on social media platforms have gained tremendous attention among researchers. Fake news through online platforms poses unique challenges. Firstly, it is written intentionally and is subjective, making it very difficult to authenticate it based on news content. Secondly, social media information is unstructured and multi-modal, both aspects are complex to capture and integrate in fake news detection. Thirdly, fake information spreads very quickly and is mainly circulated through bots, trolls, and humans from varied backgrounds. Identifying such spreaders and victims is a challenging task. This article presents a critical review of the literature on fake news detection. The state-of-the-art methods are discussed, most of the methods depend on news contents, user profiles, and social context features of a post. The importance of feature engineering, feature extraction and feature fusion in FND are highlighted. Various fake news detection datasets are discussed. Finally, future research directions in the fake news detection problem are presented.

Keywords: Fake News Detection, Textual Features, Visual Features, social media, News Contents, Social Context, Machine Learning, Deep Learning.

1. Introduction

In the internet era, social media platforms such as Facebook, Twitter, WhatsApp has become an integral part of people's lives. They help people to stay connected, updated about trending events and are the primary means of sharing news and user opinions, but these platforms have serious side effects as well. These platforms have now become a means of spreading fake and unverified information. Misleading information appearing as textual news (headline and body), photo shopped images, doctored videos remains a concern. The news articles with images, videos are more appealing and attract more attention from readers than the traditional newspapers.

Social media platforms use persuasive technology to keep users engaged and increase their screen time, strong recommendation helps people connect to like-minded people and groups. Social media popularity indicators, "likes" on Facebook, "thumb up" for YouTube videos, et al. contribute in deciding the authenticity of the message. People believe in whatever they see. Research shows that, psychologically, if a post on social media has more likes and comments, it can change other readers' perception towards the quality of the message and views about the topics discussed in the message. In additional social dynamic from popularity, indicators says: "When a post is accompanied by many likes, shares, or comments, it is more likely to receive attention by others, and therefore more likely to be further liked, shared, or commented on" \footnote{How is Fake News Spread? Bots, People like you, Trolls, and Micro targeting [Online].Website https://www.cits.ucsb.edu/fake-news/spread [accessed: 10 January 2022].}

The Internet and social media are easily accessible to everyone. With little verification process, people can create websites, blogs, and social networks platform accounts. Through this, a huge volume of fake content is published and shared every day. Such websites and accounts are fake their main aim is to circulate hoaxes, propaganda messages mostly related to politics and finance. The majority of fake news sharing happens by humans (real human accounts) knowingly or unknowingly. Since people like novelty and fake news describes events that are unique and never happened. The propagation of fake news happens much faster as compared to real news. Such social interactions in the form of discussions, comments, likes and dislikes, are called social context features of news posts.

Fake news propagation has become a worldwide concern today. It can influence the well-being of nations. The growing fake news problem has prompted The Prime Minister of India Narendra Modi to address the same in the NAM summit\footnote {PM Modi at NAM Summit: terrorism, fake news "deadly virus" [Online].Website

https://indianexpress.com/article/india/pm-modi-at-nam-summit-terrorism-fake-news-deadly-viruses-6394202/ [accessed: 10 February 2022].}.

Detecting fake or misleading content on social media poses a unique challenge. Firstly, fake news is subjective and depends on the topic or event under discussion. Secondly, fake news mimics real news in terms of writing style. Most of the time it is syntactically and semantically correct but untrue. It is deliberately written to mislead readers. Thirdly, multiple modalities are considered while creating fake news. For example, a social media post can be made up of any combination of text, images, audio, video, infographics, et al. Finally, news on social media is constantly updating making it difficult to verify it against the available knowledge base (Agarwal, & Dixit, 2020).

The main contributions of this paper are outlined as follows:

- The Characteristics of fake news on social media platforms are identified and discussed. These characteristics play a major role in deciding whether a post is real or fake.
- Major feature extraction and feature reduction techniques used in literature for textual and visual data are discussed.
- An in-depth review of single modality-based machine learning and advanced deep learning techniques for fake news detection is provided.
- The need for multi-modal fake news detection systems is highlighted. A detailed review of multi-modal-based advanced deep learning with more emphasis on feature fusion is presented.
- The publicly available datasets used in literature for FND problem are presented.
- Future research directions are briefly outlined.

The remainder of this paper is organized as follows. Section II gives a problem definition. Section III presents the definition, components of fake news and discusses various feature engineering techniques with respect to textual, visual, and social context features. Section IV gives a review of prior work on fake news detection. Section V provides details about the currently available datasets for the FND problem. Section VI presents open issues and future research directions. Section VII provides the conclusion of the work.

2. Problem Statement

We consider fake news detection as a binary classification problem i.e., classifying a social media news post as real or fake. A news post on social media consists of text, visual content like images and video, and social context information such as likes, shares, comments, et al. Let P be a collection of such news posts on social media.

$$P = \{(M_1, S_1), (M_2, S_2), (M_3, S_3) \cdots (M_N, S_N)\} \ldots\ldots\ldots\ldots\ldots(1)$$

where M_i →i^th social media post with text and visual (image or video) information.

S_i →social context information of i^th post.

N → total number of posts.

Consider a single i^th post on social media. This post 'i' has textual features T_i generated from text , visual features captured from attached image or video denoted by V_i and social context features denoted by S_i. A conceptual representation of these features is given below.

$$T_i = \{T_i^1, T_i^2, T_i^3 \cdots T_i^r\} \ldots\ldots\ldots\ldots\ldots(2)$$

$$V_i = \{V_i^1, V_i^2, V_i^3 \cdots V_i^q\} \ldots\ldots\ldots\ldots\ldots(3)$$

$$S_i = \{S_i^1, S_i^2, S_i^3 \cdots S_i^s\} \ldots\ldots\ldots\ldots\ldots(4)$$

Where r, q, s depends on the deep learning architecture under consideration.

The aim is to design a model which takes a post as input, generates its textual, visual, and social context features, and classifies it to a predefined label as a real post or fake post.

$$\theta: p \Rightarrow \{T_p, V_p, S_p\} \Rightarrow O \ \ldots\ldots\ldots\ldots\ldots(5)$$

where O is a predefined label, $O \epsilon \{real, fake\}$, θ is the learned model.

3. Feature Engineering in Fake News

Before performing feature engineering and extracting the relevant features, it's important to understand the fake news and its major components.

3.1. Fake news and its components: The widely adopted fake news definition is "Fake news is a news story written with dishonest intentions to mislead the readers and contains verifiable false information" (Shu, K. & Liu, H. (2019). The definition is based on two key features: authenticity and intent. Authenticity talks about the contents of fake news which are verifiable and intentionally tells the deceptive motive of the creator of fake news (Shu, K. & Liu, H. (2019). Fake news has four major components: news contents, social context, creator, dissemination/spreader. Figure 1 explains in detail every component and its features.

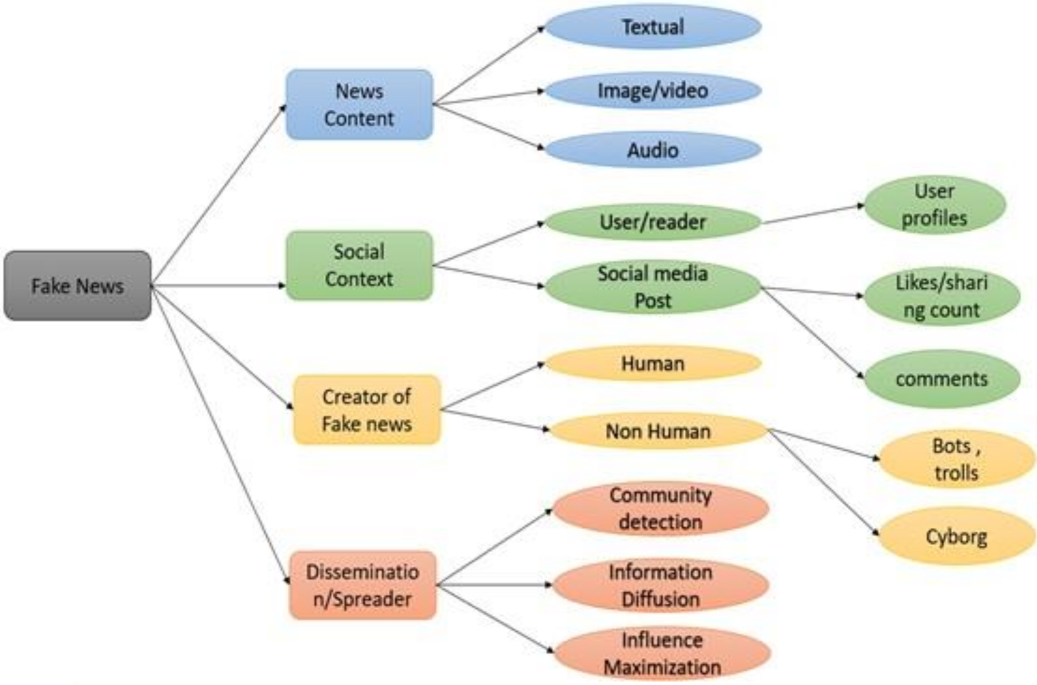

Figure 1. Fake News Characteristics

1. News Content: It refers to the news headline, body of news, and supporting images, videos, or audio. News content is multi-modal in nature i.e., a piece of news can be text only, image/video only or audio only, or a combination of text, images, videos and audio. Every component of the news like URL (Uniform Resource Locator), hashtags, and mentions are important \cite{ref10} and are considered as news content.

2. Social context: It refers to factors that play a major role in the dissemination of news on social media platforms \cite {ref10}. It includes discussion, comments given by users, likes, shares, and retweets on news posts. It provides valuable insights regarding news articles being authentic or hoaxes. It also helps to recognize the distribution pattern of real and fake news stories (Zhang, X. & Ghorbani, A., 2020).

3. Creator of Fake News: Creator of fake news can be humans or non-humans like bots (Zhang, X. & Ghorbani, A., 2020). Humans are malicious users who create fake news for a purpose. Bots are computer programs that mostly help in spreading fake news.

4. Dissemination/spreader: The creator of fake news always decides the people/users who will be influenced by fake news and will maximum participate in its spread. The Spread of fake news along with humans is also done by bots or cyborgs. Fake news dissemination can be done by identifying the target communities, understanding information diffusion, and identifying targets where influence maximization can be achieved.

Valuable information in the form of features can be extracted from the fake news components. These extracted features can be further analyzed and can be used by predictive models. Various techniques to extract textual, visual, and social context features are discussed below.

3.2. Textual feature extraction Techniques: The information obtained from the news title and news body are textual features. They are classified into four types i.e. semantic, syntax, lexicon, and discourse (Shah, P., 2020).

Syntax features: These are sentence-level features and can be captured using Bag-of-word (BoW)s, n-gram, and Parts-of-speech (POS) tagging and Context-Free Grammar (CFG) analysis[16,17].

Lexical features: It is used to capture the character and word-level information. It gives the statistics of words and letters in the text-gram models [16, 17].

Semantic and psycholinguistic features: Semantic features help to understand the meaning of data and psycholinguistic features helps to capture persuasive and biased language. Google's API (https://www.perspectiveapi.com) and Linguistic Inquiry and Word Count (LIWC) are used to extract these features [16, 17].

Feature selection and feature reduction techniques are used to lower the dimensionality of textual features. Feature selection is a process that selects a subset of relevant features from the original feature set. The feature selection methods are Filter, Wrappers, and Embedded Methods. Chi-square test, Document Frequency (DF), Information Gain (IG), Best Term (BT), Ambiguity Measure (AM), and Distinguishing Feature Selector (DFS) are commonly used filter methods of feature selection. Feature reduction helps to get a new set of features from features at the feature selection stage. Principle Component Analysis (PCA) and Latent Semantic Indexing (LSI) are commonly used feature reduction methods (Haylat, T., 2020).

(Dzisevič, R. & Šešok, D., 2019) captured text features using three different feature extractors and highlighted the one that allows the classifier to give the best accuracy. They used Term Frequency- Inverse Document Frequency (TF-IDF) and its two variations namely TF-IDF with LSA (Latent Semantic Analysis) and TF-IDF with LDA (Linear Discriminant Analysis). (Bharadwaj, P., Shao, Z., & Darren, S., 2019) extracted semantic features using TF-IDF, unigram, bigram, N-gram, and recurrent neural network (RNN). They further highlighted that semantic features can be combined with linguistic clues and metadata to improve detection.

3.3. Visual Image Features Extraction: Due to advancements in technology, many user-friendly sophisticated image editing tools are available in the market. Because of these distinguishing between tampered and real images through the naked eye becomes difficult. Such images are used in immoral ways such as adding it to a misleading or fake news post et al. (Abidin, Majid, Samah, & Hashim, 2019). Mostly used techniques for digital image forgery are copy-move, splicing, morphing, resampling, and compression [15, 16].

Discrete Cosine Transform (DCT) based methods, block feature extraction using Fourier -Mellin transforms (FMT), Discrete Wavelet Transformation (DWT) using pixel matching, Speed up Robust Feature (SURF), Scale-invariant feature transform (SIFT) based methods are used for identifying copy-move forgery. Color Filter Array (CFA), Discrete Octonion cosine transformation (DOCT), and histogram techniques are used to identify image splicing [15, 16]. Many machine learning classifiers in combination with image features are also used for image forgery detection. However, the limitation of this method is that they are suitable for an individual forgery type. When multiple tampering is applied over a single image, the accuracy of algorithms will start decreasing (Singh, B. & Sharma, D., (2021).

Deep learning algorithms can extract important features on their own whereas machine learning algorithms require explicit feature engineering. Deep learning also performs better when there are multiple manipulations in the image and it can learn the forged image features without explicit help from the training dataset (Singh, B. & Sharma, D., 2021). Related research presented then section 5 explains in-depth various deep learning approaches used for image feature extraction.

3.4. Social Context feature extraction: Social context refers to the news propagation on social media. Social Context features include user-news engagements, user and its friend's network information, count of likes and dislikes for a news article. (Hlaing & Kham, 2020) described the process of collecting the social context features. Application programming APIs provided by social media platforms were used to collect user-news engagements such as likes, dislikes, comments, reposts, et al. APIs are also used to collect metadata, user profiles, social network information.

4. Related Research in Fake News Detection

There is vast research currently going on in fake news detection (FND) on social media using Artificial Intelligence-based techniques.

4.1. Single-modality-based Fake News Detection

4.1.1. Deep Learning Approaches: (Girgis, Amer, & Gadallah, 2018) implemented a deep learning model considering online textual news content. They namely used the vanilla, GRU (Gated Recurrent Unit), and LSTM (Long Short Term Memory) model on the LAIR dataset. This model provided better results than the traditional linguistic clues approach. Vanilla was not suitable for practical problems. Compared to GRU, LSTM was inefficient as it is more expensive to calculate network output. The best results were given by GRU as it solved the vanishing gradient problem. They mentioned that CNN combined with GRU can give more accurate results (Girgis, Amer, & Gadallah, 2018). (Ajao, Bhowmik, & Zargari, 2018) propose a system where given a tweet about a news item, the system will determine whether it is true or fake based on the content of the message. They aim to identify the linguistic characteristic linked to the news automatically without prior domain expertise, through the hybrid CNN and LSTM model (Ajao, Bhowmik, & Zargari, 2018). It takes into account only Twitter posts and is not able to track the geo location and origin of fake news. (Bharadwaj, P., Shao, Z., & Darren, S., 2019) extracted semantic features from news posts using TF-IDF, unigram, bigram, N-gram, recurrent neural network (RNN), Naïve Bayes(NB), and random forest are used for further classification. They further highlighted that semantic features can be combined with linguistic clues and metadata to improve detection. (Dong, et al., 2019) used attention forest for detecting opinion and fact-based false information. They used attentive bidirectional GRU for textual feature extraction and a deep neural network for extracting features for side information. To assess news credibility on social media, (Kaliyar, Kumar, et al., 2020) created a deep neural network. Along with news content, user profiles and user groups are taken into account. The news-user engagement and user community information are combined into a 3-D tensor. A tensor factorization method is also employed, yielding a latent design of both news content and social context. Artificial Neural Networks (ANN) and Deep Hybrid Neural Networks (DHNN) were the categorization models used. To enhance the accuracy of fake news identification, the authors plan to integrate temporal information about the dissemination of fake news (Kaliyar, Kumar, et al., 2020).

4.1.2 Ensemble-based approaches: In (Agarwal, & Dixit, 2020), authors presented an ensemble learning approach for addressing the problem of fake news. The ensemble classifier was created using SVM (Support Vector Machines), convoluted neural network (CNN), LSTM, KNN(K-Nearest Neighbour), and NB as basic classifiers. Linguistic features are extracted from the news. Extracted features are correlated with the author of the news article and the credit score is calculated. It was observed that authors with higher credit scores are less likely to form fake news. (Kaliyar, Goswami, & Narang, 2019) developed a multi-class tree-based ensemble classifier using gradient boosting with optimized parameters. TF-IDF, Cosine Similarity, Hand Selected Features, Word Overlap Features, Polarity Features, and Refuting Features are used to extract content and context features from news articles. In the future, the authors plan to apply an optimized deep learning model and a powerful language model like BERT (Kaliyar, Goswami, & Narang, 02019).

4.1.3 Text and social context-based approaches: The majority of current FND algorithms are focused on news content, which is less effective because false news is intentionally created to deceive readers by imitating actual news. Thus, news content should be combined with some supplementary information to enhance detection. (Shu, Wang, & Liu, 2019) used social context information such as user credibility, and publication credibility along with news content. They created TriFN, which considers both publisher-news and user-news relations at the same time for fake news classification. They highlighted that psychology's perspective of the creator and malicious user spreading fake news should be identified for effective fake news intervention and mitigation (Shu, Wang, & Liu, 2018). demonstrated that a correlation exists between user account profiles and fake/real news spread on social media. Users' likelihood of believing fake news has different characteristics than those believing real news. Comparative analysis considering the explicit and implicit user profile features was presented. FakeNewsNet is a data repository of fake news articles created by (Shu, Mahudeswaran, Wang, Lee, & Liu., 2020). It offers two comprehensive data sets, for every new article in the dataset along with news content, spatiotemporal, and social context information is also provided. News content features were taken from fact-checking websites, social context features obtained from Twitter's advanced search API, and spatiotemporal information extracted from user profiles. They further highlighted that the FakeNewsNet repository can be integrated with front-end software and build an end-to-end system for fake news study.

(Hlaing & Kham, 2020) used social context features along with news content for FND. They collected social context information like reaction counts, comments, and content from Facebook using graph API and legitimate news stories from News API. The semantic similarity match between Facebook posts and legitimate news stories was done using WordNet. The polarity of comments was calculated using VADER. Finally, Adaboost, Decision tree, and Random Forest classifier were used to calculate news authenticity score. Authors highlighted that considering social context features along with news content for FND is a challenging task and a multi-dimensional benchmark dataset is necessary for further research (Hlaing & Kham, 2020).

4.1.4 Single Modal Visual Feature Detection: Deep learning techniques have self-feature extraction capability (Majumder, M.T.H. & Alim Al Islam, A.B.M., 2018) and with the power of GPU (Abidin, Majid, Samah, & Hashim, 2019). they provide better performance than conventional, machine learning approaches that require domain expertise. But the drawback of deep learning is that large datasets are required for training and validation (Abidin, Majid, Samah, & Hashim, 2019). In literature, mostly CNN models are used for image feature extraction. (Majumder, M.T.H. & Alim Al Islam, A.B.M., 2018) proposed the use of a shallow CNN for image forgery detection. (Kaliyar, Goswami, & Narang, 02019) used a pre-trained AlexNet model for copy-move forgery detection in images. (Singh, B. & Sharma, D., 2021). used 16 high-pass filters to amplify the noise in the image, then CNN is used to learn the intrinsic features of the image. The gradient information of the last

DOI: 10.6977/IJoSI.202306_7(6).0004
D. Nikumbh, A. Thakare/Int. J. Systematic Innovation, 7(6), 36-53(2023)

Table 1: The Summary of Single Modal Fake News Detection Approaches

Reference no	Textual feature extraction	Social context features	Creator features	Datasets	Classifier	Accuracy	Future scope
8	RNN: Vanilla, LSTM and GRU	No	No	LAIR	-	-	Combine GRU with CNN.
7	LSTM+CNN	No	No	PHEME	-	LSTM:82.29% ,LSTM +dropout: 73.38%, LSTM-CNN:80.38%	Tracking the origin and location of fake news.
12	TF-IDF, N-gram, Glove, RNN.	No	No	*real-or-fake news* dataset from kaggle.com	Naïve Bayes and Random Forest.	Bi-grams with random forest: 95.66%	Combine semantic features with linguistic clues and metadata.
35	Word2Vec, POS tagging	No	Yes	Combines LAIR and dataset from Kaggle.	Ensemble of classifier(SVM , CNN, LSTM, KNN, and NB)	Ensemble classifier: 85%	-
21	Attentive Bidirectional Gated Recurrent Unit(GRU)	Yes	No	Politifact, Facebook Fact-check, and Amazon review dataset	Attentive Forest	AttForest-C: Politifact:80.40% Factcheck:83.30% Amazon:94.80% Attforest2: Politifact:82.80% Factcheck:84.40% Amazon:96.70%	Include more clues like images and videos.
22	TF-IDF features, Cosine Similarity Features, Hand Selected Features, Word Overlap Features, Polarity Features, Refuting Features	No	No	Fake News Corpus(FNC)dataset	Tree-based ensemble classifier	86%	Apply optimized deep learning models and powerful word techniques like BERT

36	Semantic Similarity using WordNet.	Yes	No	Own dataset using Graph API, News API, and referring BuzzFace dataset.	Decision Tree, Adaboost, RandomForest	The accuracy of random Forest is better than the other two classifiers.	Use of hybrid classification methods to improve performance.
26	Clauset-Newman-Moore algorithm	No	Yes	BuzzFeed and Fakeddit	Artificial Neural Network(ANN), Deep Hybrid Neural Network(DeepNet)	ANN:82% DeepNet: 86.40%	Include the temporal information.

convolutional layers was used to localize the target region of manipulation in the image.

More research is needed in image forgery detection for real-world images, multiple tampered images, homogenous images (Abidin, Majid, Samah, & Hashim, 2019), and fake images generated by GAN (Generative Adversarial Network) (Singh, B. & Sharma, D., 2021). For deep learning models, transfer learning and using different learning rates at different layers should be explored to increase accuracy (Girgis, Amer, & Gadallah, 2018).

4.2 Multi-modality-based Fake news detection techniques: Single-mode techniques produce promising results, but the majority of the content on social media platforms nowadays is unstructured. (along with text there can be images, audio, or video). Researchers are now focusing on extracting from such unstructured multi-modal data. Various models are developed which consider both text and image for FND. (Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., Su, L., & Gao, J. (2018) developed an architecture to extract event invariant features from multi-modal posts using an adversarial technique. They highlighted that existing approaches extract event-specific features from news posts, which is ineffective for detecting fake news for new events. The minimax game is set between a multi-modal feature extractor and an event discriminator. Through this approach event invariants features are learned, which are given to fake news detectors to classify the post as real or fake. (Singhal, Shah, Chakraborty, Kumaraguru, & Satoh. (2019) highlight that current fake news detectors perform sub-task like event discriminators. If the subtask training is not performed it can degrade the performance of the model. They used BERT a language model to capture both context and content features and pre-trained VGG-19 for image features. A simple concatenation approach is used for fusing textual and visual features. The authors highlighted that the study of different modalities in fake news detection and more complex multi-modal fusion techniques must be explored (Singhal, Shah, Chakraborty, Kumaraguru, & Satoh., 2019).

(Khattar, Goud, et al., 2019) used variational autoencoder approach to learn combined multi-modal features representation. LSTM's are used in autoencoders for extracting textual and visual features. The authors plan to extend the model to social context features like tweet propagation and user characteristics (Khattar, Goud, et al., 2019). (Zhang, et al., (2020). developed a BERT-based model with a domain classifier. The domain classifier is responsible for removing event-specific dependencies from multi-modal features. Authors further intend to use the proposed model on similar other datasets, develop a probabilistic FND model, and indicate the relevance of the attached image to text in the post while performing fake news detection (Zhang, et al., 2020). (Tanwar & Sharma, 2020) also used a variational encoder to obtain a shared representation of multi-modal features. Here in the encoder three CNN architectures namely Inception V3, ResNet 50, and VGG-19 are used for image feature extraction. Authors further want to test their model on other publicly available datasets and incorporate additional features like user profile data to enhance accuracy (Tanwar & Sharma, 2020). (Madhusudhan, Mahurkar, & Nagarajan, 2020). used two different multi-modal fusion methods for textual and visual features. In one method textual and visual features are extracted independently and concatenated and in the second approach visual attention is applied. For extracting textual features BERT and SBERT were used and for image features pre-trained ResNet18 was used (Madhusudhan, Mahurkar, & Nagarajan, 2020).

(Giachanou, Zhang, & Rosso, 2020) developed a multi-modal multi-image FND system. The model along with the textual and visual information, uses semantic information as well. The novelty of the work is that multiple images of posts are considered for extracting visual features, temporal information among the images is captured by LSTM (Long Short-Term Memory) and a similarity score is computed among the text and image tags. All three features i.e. text, visual, and semantic are fused either through concatenation or through attention mechanism for making the predictions. (Shah, P., 2020) used an evolutionary computing approach for fake news detection. For a multi-modal news article, the author extracted textual features using sentimental analysis and image features using thresholding and segmentation. A cultural algorithm is used for optimizing the textual and visual features extracted from a news article. She further used an SVM classifier on optimized features (Shah, P., 2020). The author wants to extend the research by considering user-independent features like demographic, sex, age, and reading pattern of readers, social media post dissemination pattern in the model.

Table 2: The Summary of Multi-Modal Online Fake News Detection Approaches

Reference no	Textual feature extraction	Visual feature extraction	Social context features	Creator features	Datasets	Accuracy	Future scope
2	BERT	VGG-19	No	No	Twitter and Weibo	Twitter:83% Weibo:86.50 %	Used developed model on similar fake news dataset. Plan to employ a probabilistic model Give relevance of the attached image to text while performing fake news detection.
1	Word2Vec +bidirectional LSTM	VGG19+ResNet50+InceptionV3	No	No	Twitter	Twitter:76%	Use the developed model on other publicly available dataset. Consider features like user profile to increase accuracy of model.
27	BERT, SBERT	ResNet pre-trained in ImageNet.	No	No	Gossipcop Politifact	Gossipcop: Baseline+BERT+Text=89.90% Visual attention+BERT+Text=89.80% Politifact: Baseline+SBERT+Text=89.70 %	-
28	BERT	VGG16+LSTM	NO	NO	Created own dataset and used a part of FakeNews Net dataset.	3-image-vgg16-LSTM+BERT+similarity+fusion(attention)=79.55 %	-

13	Sentimental Analysis	Segmentation using KNN and DWT(Discrete Wavelet Transformation)	No	No	Twitter, Weibo	Twitter: 79.80% Weibo: 89.10%	Consider user independent features and social media post dissemination pattern in model.
3	Bi-directional LSTM's	VGG-19	No	No	Twitter ,Weibo	Twitter: 74.50% Weibo: 82.40%	Extend the model considering tweet propagation and user characteristic.
4	BERT Base	VGG-19	No	No	Twitter ,Weibo	Twitter: 77.77% Weibo: 89.23%	Explore more complex fusion techniques and how different modalities important in fake news detection.
5	Text-CNN	VGG-19	No	No	Twitter ,Weibo	Twitter: 71.50% Weibo: 82.70%	-

4.3 Multimodal Feature Fusion: While building a multimodal fake news detection system, it's very important to focus on fusion techniques. Fusion techniques bring information from different modalities together. Most of the models developed in literature take linear combinations or simple concatenation of modalities whereas complex interaction between the modalities should be explored to develop efficient models.

Figure 2 shows different feature fusion approaches implemented in literature. The early fusion approach is also called data-level fusion. It is applied to raw data or pre-processed data. Here feature extraction of independent modalities is done followed by feature fusion which results in a single feature vector. Early fusion assumes conditional independence between multiple modalities, which is not always true like in the case of video and depth clues (Haylat, T., 2020). The simplest form of early fusion is the concatenation of extracted multimodal features into a single shared representation. Dimension reduction techniques like Independent Component Analysis (ICA), PCA, and canonical correlation can be applied to the extracted features, to make them dimensionally identical, as this will facilitate concatenation operation. Many papers in the literature son multi-modal fake news detection implement early fusion techniques. (Leyva, R. et al., 2019) used an early fusion mechanism to fuse the text, image, and audio features in video memorability prediction systems. The features extracted from different modalities were first subjected to feature reduction using PCA(Principal Component Analysis). Each of these reduced features was stacked and finally given to the regression model to calculate the memorability score.

DOI: 10.6977/IJoSI.202306_7(6).0004
D. Nikumbh, A. Thakare/Int. J. Systematic Innovation, 7(6), 36-53(2023)

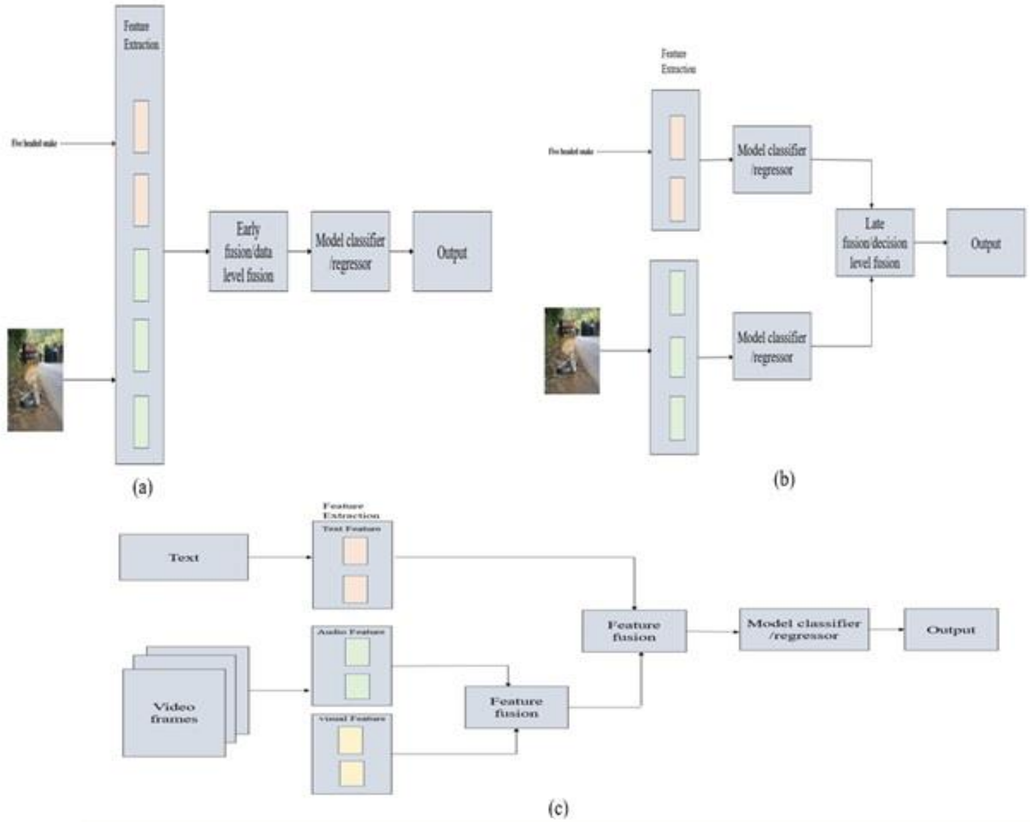

Figure 2. Multimodal feature fusion Techniques. (a) Early fusion approach, (b) Late fusion approach, and (c) Intermediate fusion approach.

The late fusion approach is also called decision-level fusion and is inspired by ensemble classifiers. If the multimodalities are uncorrelated in terms of data dimension, a unit of measure, late fusion approaches give better results. Different rules like the Bayes rule, maximum rule fusion, and average fusion are some of the late fusion techniques (Haylat, T., 2020).

(Kampman, O., Jebalbarezi, E., Bertero, D. & Fung, P., 2018) developed an automatic personality prediction where they investigated different fusion methods for three channels namely audio, text, and video. The first fusion methods are decision-level fusion or late fusion, implemented through the ensemble(voting) method. In this technique, an estimator score was calculated for each personality trait. The estimator score is a weighted sum of the estimator for each trait considering each modality. The advantage of this method is the relevance of modality for a particular personality trait can be identified from the weights (Kampman, O., Jebalbarezi, E., Bertero, D. & Fung, P., 2018). The second fusion method uses limited backpropagation, where only the last two layers of architecture are trainable and the third fusion method uses full backpropagation for the entire architecture (Kampman, O., Jebalbarezi, E., Bertero, D. & Fung, P., 2018).

The intermediate fusion approach is based on a deep neural network. In this approach, a single hidden layer is used to learn the combined representation of different modalities' features. This single hidden layer could be a fully connected, 2D, or 3D convolutional layer. The layer where different modalities fuse is called a fusion layer (Haylat, T., 2020). Figure 2c shows intermediate fusion where the first audio and visual features of the video frame are fused followed by the fusion of text features.

***4.3.1 Attention-based fusion techniques**:* The attention mechanism is inspired by cognitive processes where humans concentrate on particular things and ignore rest. For example, if asked to look for a cat in images, our brains find an object with cat-like features and ignore the rest. The human brain is tuned to the attention. In deep learning, an attention mechanism could be visual attention that concentrates on important regions in an image and text attention that focuses on important words in the text. Neural Machine Translation Systems were the first to implement an Attention mechanism, to overcome the long-range dependency problem of LSTM's and RNN's.

Attention mechanisms are broadly classified as self-attention and cross-attention. (Vaswani, A., et al., 2017). introduced the concept of the self-attention mechanism through transformer architecture. In the self-attention mechanism while processing an element of a sequence, which other parts of the same sequence are important to process the element is found out. The self-attention mechanism extracts intra-modality information, where query, key, and value belong to the same modality. Cross attention mechanism generates inter-modality information, where for example query can be based on text input and key and value based on image input.

(Duc Tuan & Quang, 2021) addressed the issue of fusing multi-modal features through cross and self-attention mechanisms. Cross attention is used to represent a correlation between text and image and vice versa of a post. Self-attention is used to represent a correlation between different image regions of an image in a post. Further, a scaled dot product attention is used to fuse text and images feature. The formulas for generating query, key, value, and cross attention by (Duc Tuan & Quang, 2021) are as follows:

$$Query(Q) = T_f \times W_Q$$
$$Key(K) = I_f \times W_K$$
$$Value(V) = I_f \times W_V$$

where T_f is text feature vector, I_f is image region feature vector, W_Q, W_K, W_v are weight matrices and $\times$ denotes matrix multiplication operation.

The scaled dot operation is applied on K,V,Q to calculate final attention.

$$attention(text \rightarrow image) = softmax(\frac{Q \times K^T}{\sqrt{d}}) \times V \text{ ……………..(A)}$$

where $\frac{1}{\sqrt{d}}$ is scaling factor

Similarly, for $attention(image \rightarrow text)$ is calculated using the same formula where a query is formed using image feature and key and value are formed using text features.

The Q, K, V for image self-attention are generated as follows:

$$Query(Q) = I_f \times W_Q$$
$$Key(K) = I_f \times W_K$$
$$Value(V) = I_f \times W_V$$

In (Ying, Yu, Wang, Ji, & Qian, 2021b), multi-modal cross attention is implemented by first concatenation of text and visual features. The dimension of visual features is converted to the same dimension as text features. The concatenated feature vector S is then fed to the transformer to generate attention. The following formulas are used in (Ying, Yu, Wang, Ji, & Qian, 2021b) for generating the query, key, and value.

$$S = \begin{bmatrix} T_f \\ I_m \end{bmatrix}$$
$$Query(Q) = S \times W_Q = \begin{pmatrix} T_f W_Q \\ I_m W_Q \end{pmatrix} = \begin{pmatrix} Q_{Tf} \\ Q_{Im} \end{pmatrix}$$
$$Key(K) = S \times W_K = \begin{pmatrix} T_f W_K \\ I_m W_K \end{pmatrix} = \begin{pmatrix} K_{Tf} \\ K_{Im} \end{pmatrix}$$
$$Value(V) = S \times W_V = \begin{pmatrix} T_f W_V \\ I_m W_V \end{pmatrix} = \begin{pmatrix} V_{Tf} \\ V_{Im} \end{pmatrix}$$

The attention is calculated by using SoftMax as mentioned in equation (A).

(Wang, Mao, & Li, 2022) designed a fine-grained fusion model using a scaled dot product mechanism. Several scaled dot product attention blocks are applied to enhance the textual and visual features. The enhanced features are further passed to two more blocks which perform inter-attention. The output is refined features which contain a fusion of textual and visual features. Authors further plan to fuse social context features in addition to textual and visual features. Visual features in frequency domain are to be considered for improvement in the model.

(Ying, Yu, Wang, Ji, & Qian, 2021a) highlighted that existing work suffers from low generalization if a post is related to a very rare or new topic. Hence topic modeling is crucial and should be integrated into fake news detection models. Along with attention (inter and intra modality) to capture post representation they also incorporated a topic memory network to capture global topic features.

(Xue, et al., 2021) designed a unique model which consists of five subnetworks namely text feature extractor, visual feature extractor, tampered visual feature extractor, similarity and fusion modules for FND. The similarity module obtains a semantic representation of text and visual features and cosine similarity is used to measure similarity between them. An attention mechanism is used in the fusion module to assign weights to physical (tampered features) and semantic features. The authors further proposed to perform optimization at the feature fusion level.

(Liao, Q., et al., (2021) designed a model for short fake news detection through multi-task learning. The authors proposed a novel N-Graph method that learns textual and contextual relations in news simultaneously in the representation learning phase. Multi-task learning module simultaneously performs FND classification and news topic classification, a dynamic weight strategy is incorporated during multi-task learning.

(Kumari, & Ekbal, 2021) introduced a novel feature fusion technique named Multimodal Factorized Bilinear Pooling. The authors further argued that semantic alignment between text and images should be investigated and a better fusion mechanism should be designed. Videos modality should also be considered in FND.

Most of the existing research focuses on implementing supervised learning approaches for FND whereas work on unsupervised and semi-supervised approaches is scarce. (Li, Guo, Wang, & Zheng, 2021). developed an unsupervised FND model using autoencoders. The model considered text, image features, news propagation, and user features on social networks. A splicing method is used to fuse the multimodal features. The author further intends to include more social context features like comments, dissemination patterns of fake news, and other modalities like videos. Also, a detailed classification model should be developed.

(Dong, Victor, & Qian, 2020) used a semi-supervised learning approach for FND. They developed a model with three CNNs: Shared CNN is used to learn low-level features, which are further passed to supervised and unsupervised CNN respectively. For calculating the loss of supervised path cross entropy measure is used and for calculating loss of unsupervised path MSE measure is used. The final loss is optimized using Adam optimizer. The author intends to use the proposed model for various other NLP tasks like sentiment analysis and dependency tasks.

Feature fusion is an important aspect of multimodal fake news detection. Various techniques like attention-based fusion mechanisms should be explored, as they provide the relation between text and supporting images of a post, which is very helpful in detecting misleading posts on social media.

5. Datasets

The publicly available datasets, used in literature for fake news detection(FND) problem are as follows:

5.1 LAIR: It is a publicly available dataset published in 2017 (Wang, W. Y., 2017). It contains short statements related to politics and is extracted from politifact.com. The dataset contains 12,836 samples. For every sample short statement, speaker, context, label, and justification fields are provided. LAIR is a multiclass dataset. Every data sample has one of 6 labels i.e true, false, pants-fire, mostly true, barely true, and half true.

5.2 Fake News Corpus -1(FNC-1): It is a news dataset that maintains news headlines and news body. It consists of 75,385 samples, every sample is labeled with one of the following labels: unrelated, agree, disagree and, discuss\footnote{Stance Detection Dataset for FNC-1 [Online]. Website https://github.com/FakeNewsChallenge/fnc-1 [accessed 12 February 2022].}. This dataset is mostly used for stance detection.

5.3 BuzzFeed: It consists of approximately 2000 news samples which are collected from Facebook during October 2016. These news articles are verified by journalists of BuzzFeed. The labels provided are mostly false, no factual content, mostly true, a mixture of true and false \footnote{Fact-Checking Facebook Politics Pages. github [Online].Website https://github.com/BuzzFeedNews/2016-10-facebook-fact-check [accessed 12 February 2022]}.

5.4 CASIA: It is a dataset of tampered images. The images are tampered with using crop and paste method. It contains 4795 images, 1701 authentic and 3274 forged (Dong, Wang, & Tan, 2013). Casia v2.0 is also available (Zheng, Y., 2019).

5.5 Twitter India Dataset: It is a dataset that contains images circulated on Twitter in India during the period November 2019 to November 2020. It covers events related to politics and religion. It has a total of 110 images out of which 61 images are fake \footnote{Twitter India Dataset Version 3. github [Online]. Website https://github.com/bhuvaneshsingh80/Twitter_India_dataset_Ver_03}.

5.6 Fake News Net: It contains news articles collected from Politifact and gossipcop. For each data sample, it maintains the following information: unique id, publisher URL, the title of the news article and tweeter id's sharing the news \footnote {FakeNewsNet. github [Online]. Website https://github.com/KaiDMML/FakeNewsNet https://github.com/KaiDMML/FakeNewsNet.}.

5.7 Twitter: It is a dataset of tweets provided by MediaEval benchmark used for identifying fake information on Twitter. The dataset contains tweet text, attached images/videos, and social context information of tweets. It contains 7898 fake news tweets and 6026 real news tweets and 514 images (Boididou, Andreadou, Papadopoulos, Dang-Nguyen, et al. (2015).

5.8 Weibo: This dataset contains rumors and fake messages circulated on a Chinese microblogging website called Sina Weibo collected during the period May 2012 to January 2016.These messages are verified by Weibo's official. The real news articles in this dataset are collected from an authentic news source of China and Weibo (Jin, Z., Cao, J., Guo, H., Zhang, Y. & Luo, J. (2017).

6. Open Issues and Research Directions

The detection of Fake news (FND) on social media has many open issues and research directions that require the attention of researchers. We suggest the following research directions:

6.1 Use of Hybrid models in FND mechanism: Most approaches in current literature work independently either on textual news contents, visual contents, or social context information. Hybrid models considering multimodal news contents combined with social context information can be a way forward.

6.2 Multi-Modal Feature Fusion: Studies in the area of multi-modal feature fusion are very limited (Mridha, Keya, Hamid, Monowar, & Rahman, 2021). Fusion approaches help to explore the correlation between text and visual data. This is very important in the case of fake news as text and image data when seen independently can be correct but when seen together might not make any sense.

6.3 Developing Large-Scale Multidimensional Dataset: For supervised learning models to work large-scale benchmark-labeled datasets are required. The lack of such datasets is causing bottlenecks in developing effective FND systems. A publicly available comprehensive large-scale dataset consisting of multimodal news contents, social context information, and dissemination pattern information is needed.

6.4 Unsupervised learning techniques for fake news detection: The availability of limited labeled datasets for FND problem has constrained the usage of supervised learning techniques. Hence, an alternative approach of using unsupervised or semi-supervised algorithms must be explored.

6.5 Fake news monitoring systems: Real-time visualization is an important aspect of monitoring systems. Detailed, multidimensional visualization with the help of modern tools will help to gain insights into online social information. It can help to reveal temporal-based news dissemination patterns, user unusual behavior, and facilitate human supervision.

6.6 Fake news intervention systems: Most of the work in literature fixate on developing accurate FND systems whereas fake news intervention systems are equally important. Monitoring systems can be combined with intervention systems to observe the impact of false information on users, to identify the users who are vulnerable and are easily influenced by false information and mitigation measures, and further monitoring of such users. Designing fake news intervention systems is a potential research area under fake news on social media.

6.7 Explainable AI: Since deep learning techniques involve very deep and complex architectures, it is very important to understand what's happening inside such architectures. Methods and techniques which fall under explainable AI should be considered while developing solutions in the area of fake news detection (Mridha, Keya, Hamid, Monowar, & Rahman, 2021).

7. Conclusion

Fake news or misleading content is a threat to society. Social media platforms that were once designed with good intentions are now being used for spreading false information causing distrust in society. The proliferation of such fake news can have a negative impact on society. Thus, automation of fake news detection has now become an extremely important task. This paper discusses major studies carried out in recent years to address the challenges in fake news detection. The major contributions of this paper are as follows:1)Fake news characteristics are outlined and discussed.2)In-depth knowledge of feature extraction techniques used in literature is provided.3)Exhaustive study of existing single-modal and multi-modal detection techniques.4)

Multi-modal feature fusion techniques used in literature and their importance is highlighted.5)Future research directions that researchers should consider in FND are presented.

This study will strongly help researchers to get better insights in dealing with the fake news detection problem, investigate and extend their work further and help them in building effective fake news detection and prevention tool.

References

Abidin, A.B.Z., Majid, H.B.A., Samah, A.B.A., & Hashim, H.B. (2019). Copy-Move Image Forgery Detection Using Deep Learning Methods: A Review, 2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS), 1-6. doi: 10.1109/ICRIIS48246.2019.9073569.

Agarwal, A., & Dixit, A. (2020). Fake News Detection: An Ensemble Learning Approach. 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 1178-1183. doi: 10.1109/ICICCS48265.2020.9121030.

Ajao, O., Bhowmik, D., & Zargari, S. (2018). Fake News Identification on Twitter with Hybrid CNN and RNN Models. In Proceedings of the 9th International Conference on Social Media and Society (SMSociety '18). Association for Computing Machinery, New York, NY, USA, 226–230. DOI: https://doi.org/10.1145/3217804.3217917

Bharadwaj, P., Shao, Z., & Darren, S. (2019). Fake News Detection With Semantic Features And Text Mining, Journal of Natural Language Processing, 8, 17-22.

Boididou, C. , Andreadou, K., Papadopoulos, S., Dang-Nguyen, D.-T., Boato, G., Riegler, M. & Kompatsiaris, Y. et al. (2015). Verifying Multimedia Use at MediaEval 2015. In Media Eval.

Dong, J., Wang, W., & Tan, T. (2013). CASIA Image Tampering Detection Evaluation Database, 2013 IEEE China Summit and International Conference on Signal and Information Processing, 422-426. doi: 10.1109/ChinaSIP.2013.6625374.

Dong, M., Yao, L., Wang, X., Benatallah, B., Zhang, X., & Sheng, Q.Z. (2019). Dual-stream Self-Attentive Random Forest for False Information Detection. 2019 International Joint Conference on Neural Networks (IJCNN), 1-8. doi: 10.1109/IJCNN.2019.8851765.

Dong, X., Victor, U., & Qian, L. (2020). Two-Path Deep Semisupervised Learning for Timely Fake News Detection, in IEEE Transactions on Computational Social Systems, 7(6), 1386-1398. doi: 10.1109/TCSS.2020.3027639.

Duc Tuan, N.M. & Quang Nhat Minh, P. (2021). Multimodal Fusion with BERT and Attention Mechanism for Fake News Detection. 2021 RIVF International Conference on Computing and Communication Technologies (RIVF), 1-6. doi: 10.1109/RIVF51545.2021.9642125.

Dzisevič, R. & Šešok, D. (2019). Text Classification using Different Feature Extraction Approaches, 2019 Open Conference of Electrical, Electronic and Information Sciences (eStream), 1-4. doi: 10.1109/eStream.2019.8732167.

Giachanou, A., Zhang, G., & Rosso, P. (2020). Multimodal Multi-image Fake News Detection, 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), 647-654. doi: 10.1109/DSAA49011.2020.00091.

Girgis, S., Amer, E., & Gadallah, M. (2018). Deep Learning Algorithms for Detecting Fake News in Online Text. 2018 13th International Conference on Computer Engineering and Systems (ICCES), 93-97. doi: 10.1109/ICCES.2018.8639198.

Haylat, T. (2020). Introduction To Data Fusion [Online], Website https://medium.com/haileleol-tibebu/data-fusion-78e68e65b2d1[accessed 13 February 2022].

Hlaing, M.M.M. & Kham, N.S.M. (2020). Defining News Authenticity on Social Media Using Machine Learning Approach, 2020 IEEE Conference on Computer Applications (ICCA), 2020, 1-6. doi: 10.1109/ICCA49400.2020.9022837.

Jin, Z., Cao, J., Guo, H., Zhang, Y. & Luo, J. (2017). Multimodal fusion with recurrent neural networks for rumor detection on microblogs, In Proceedings of the 2017 ACM on Multimedia Conference. ACM,

Kaliyar, R.K., Goswami, A., & Narang, P. (2019). Multiclass Fake News Detection using Ensemble Machine Learning, 2019 IEEE 9th International Conference on Advanced Computing (IACC), 103-107. doi: 10.1109/IACC48062.2019.8971579.

Kaliyar, R.K., Kumar, P., Kumar, M., Narkhede, M., Namboodiri, S., & Mishra, S. (2020). DeepNet: An Efficient Neural Network for Fake News Detection using News-User Engagements. 2020 5th International Conference on Computing, Communication and Security (ICCCS), 1-6, doi:

10.1109/ICCCS49678.2020.9277353.

Kampman, O., Jebalbarezi, E., Bertero, D. & Fung, P. (2018). Investigating Audio, Video, and Text Fusion Methods for End-to-End Automatic Personality Prediction, 606-611. 10.18653/v1/P18-2096.

Khattar, D., Goud, J.S., Gupta, M., & Varma, V. (2019). Mvae: Multimodal variational autoencoder for fake news detection. In The World Wide Web Conference, 2915–2921.

Kumari, R., & Ekbal, A. (2021). AMFB: Attention based multimodal Factorized Bilinear Pooling for multimodal Fake News Detection, Expert Systems with Applications, 184, 115412. doi:10.1016/j.eswa.2021.115412.

Leyva, R., et al. (2019). Multimodal Deep Features Fusion for Video Memorability Prediction, Media Eval.

Li, D., Guo, H., Wang, Z. & Zheng, Z. (2021). Unsupervised Fake News Detection Based on Autoencoder, in IEEE Access, 9, 29356-29365. doi: 10.1109/ACCESS.2021.3058809.

Liao, Q., et al. (2021). An Integrated Multi-Task Model for Fake News Detection, in IEEE Transactions on Knowledge and Data Engineering. doi: 10.1109/TKDE.2021.3054993.

Madhusudhan, S., Mahurkar, S., & Nagarajan, S.K. (2020). Attributional analysis of Multi-Modal Fake News Detection Models (Grand Challenge), 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM), 451-455. doi: 10.1109/BigMM50055.2020.00074.

Majumder, M.T.H. & Alim Al Islam, A.B.M. (2018). A Tale of a Deep Learning Approach to Image Forgery Detection. 2018 5th International Conference on Networking, Systems and Security (NSysS), 1-9. doi: 10.1109/NSysS.2018.8631389.

Mridha, M.F., Keya, A.J., Hamid, M.A., Monowar, M.M., & Rahman, M.S. (2021). A Comprehensive Review on Fake News Detection With Deep Learning, in IEEE Access, 9, 156151-156170. doi: 10.1109/ACCESS.2021.3129329.

Shah, P. (2020). Multimodal fake news detection using a Cultural Algorithm with situational and normative knowledge. University of Windsor.

Shu, K. & Liu, H. (2019). Detecting Fake News on Social Media, Morgan & Claypool Publishers, ISBN 9781681735825.

Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2020). Fake News Net: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media, Big Data, 8, 171-188. 10.1089/big.2020.0062.

Singh, B. & Sharma, D.K. (2021). Image Forgery over Social Media Platforms - A Deep Learning Approach for its Detection and Localization. 2021 8th International Conference on Computing for Sustainable Global Development (INDIACom), 705-709, doi: 10.1109/INDIACom51348.2021.00125.

Singhal, S., Shah, R., Chakraborty, T., Kumaraguru, P., & Satoh, S. (2019). Spot fake: A multi-modal framework for fake 71 news detection. In 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), 39–47.

Shu, K., Wang, S., & Liu, H. (2018). Understanding User Profiles on Social Media for Fake News Detection, 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), 430-435. doi: 10.1109/MIPR.2018.00092.

Shu, K., Wang, S. & Liu, H. (2019). Beyond News Contents: The Role of Social Context for Fake News Detection. In Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining (WSDM '19). Association for Computing Machinery, New York, NY, USA, 312–320. DOI:https://doi.org/10.1145/3289600.3290994

Tanwar, V. & Sharma, K. (2020). Multi-Model Fake News Detection based on Concatenation of Visual Latent Features, 2020 International Conference on Communication and Signal Processing (ICCSP), 1344-1348.doi: 10.1109/ICCSP48568.2020.9182398.

Vaswani, Ashish., Shazeer, Noam., Parmar, Niki., Uszkoreit, Jakob., Jones, Llion., Gomez, Aidan., Kaiser, Lukasz., & Polosukhin, Illia. (2017). Attention is all you need, Advances in Neural Information Processing Systems 30 (NIPS 2017)

Wang, J., Mao, H. & Li. (2022). FMFN: Fine-Grained Multimodal Fusion Networks for Fake News Detection, Applied Sciences, 12(3), 1093.

Wang, W. Y. (2017). Liar, Liar Pants on Fire: A New Benchmark Dataset for Fake News Detection, ACL Anthology.

Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., Su, L., & Gao, J. (2018). 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, 849–857.

Xue, Junxiao, et al. (2021). Detecting fake news by exploring the consistency of multimodal data, Information Processing & Management, 58(5), 102610.

Ying, L., Yu, H., Wang, J., Ji, Y. & Qian, S. (2021a). Fake News Detection via Multi-Modal Topic Memory Network, in IEEE Access, 9, 132818-132829.

Ying, L., Yu, H., Wang, J., Ji, Y. & Qian, S. (2021b). Multi-Level Multi-Modal Cross-Attention Network for Fake News Detection. in IEEE Access, 9, 132363-132373. doi: 10.1109/ACCESS.2021.3114093.

Zhang, X. & Ghorbani, A. (2020). An overview of online fake news: Characterization, detection, and discussion. Inf. Process. Manag., 57, 102025.

Zhang, T., et al. (2020). BDANN: BERT-Based Domain Adaptation Neural Network for Multi-Modal Fake News Detection, 2020 International Joint Conference on Neural Networks (IJCNN), 1-8. doi: 10.1109/IJCNN48605.2020.9206973.

Zheng, Yue. (2019). modified CASIA, IEEE Data report. doi: https://dx.doi.org/10.21227/c1h8-kf39