

A Technical Review of Fake News Detection: Progression from Textual Analysis to Multimodal and Knowledge-Aware Learning

Sriman D.¹, Pothula Sujatha²

¹Student, ²Professor, Department of Computer Science and Engineering, Pondicherry University, Puducherry, India.

E-mail: ¹srimanoffl@gmail.com, ²sujathaps.csc@pondiuni.ac.in

Abstract

The process of fake news detection has evolved from classification-based text processing to multimodal, propagation-based, knowledge-based, and heterogeneous graph learning for addressing the increasing complexity of disinformation on social media platforms. This paper presents an analysis of methodological advances made in this domain through the examination of the use of textual, visual, emotional, propagation, relational, and external knowledge data in fake news detection frameworks. The recent studies related to fake news detection have been categorized on the basis of fuzzy and multitask learning, attention and emotion modeling, multimodal fusion and consistency learning, propagation early detection, hyperbolic representation learning, knowledge-based vision-language models, heterogeneous graph neural networks, multilingual detection, and real-time analysis. This research study compares existing major datasets and evaluation practices to emphasize the impacts of modality, language, platform, temporal information, and labeling characteristics on the reported performance. The key issues related to cross-modality inconsistency, dataset biases, distribution shifts, interpretability, knowledge reliability, temporality, and computational scalability are identified and the literature-driven research directions towards robust, interpretable, multilingual, and evidence-based fake news detection are experimented.

Keywords: Multimodal Fake News Detection, Cross-Modal Consistency, Propagation-Aware Learning, Knowledge-Enhanced Vision-Language Models, Heterogeneous Graph Neural Networks, Hyperbolic Representation Learning, Evidence-Grounded Detection

1. Introduction

Rapid advancement in social media technology has fastened the creation of fake news, which is causing deceptive messages to reach more people through connectivity on digital platforms. Modern fake news is characterized by the creation of textual messages with images and other contextual details to improve their authenticity, which poses challenges in detecting fake news [3], [9]. Early methods for detecting fake news involved the linguistic and statistical feature extraction technique at first, followed by the deep learning model, fuzzy learning method, multitask learning, and self-attention techniques that improve the quality of textual representation [2], [4], [5]. However, text-based models are inefficient in scenarios where the credibility of the deception lies in their interaction.

Detection through multimodal learning is done through the analysis of both the textual and visual data using approaches like feature fusion, modal interaction, consistency modeling, and co-attention [1], [6], [8], [11] has become another development in misinformation detection due to the appearance of multimedia data. Besides, misinformation detection has been expanded to consider the role of information propagation, temporal behavior, and hierarchical representation as means for explaining the appearance and spreading of misinformation through social media [7], [14], [15]. Knowledge-enhanced vision-language model and heterogeneous graphs are two more examples of such an approach expansion [12], [13], [17].

Despite this development in methods, fake news detection is still hindered by several factors, including heterogenous data, scarcity of multimodal and multilingual sources, domain differences, problems with interpretability, and the computational issues involved in real-time processing [9], [10], [16], [18], [19]. Such challenges make direct comparison of the methods hard, and hence there is a need to examine beyond performance assessment to look at the basis on which each method works, including the data and context surrounding it. In that regard, this survey outlines the development of fake news detection from text-based detection to other types of methods.

2. Related Works

Fake news detection has evolved from text-based classification to multimodal, propagation-aware, and knowledge-based processing. The fuzzy and multitask DL methods

helped build a framework for dealing with uncertainties and learning related tasks in tandem, whereas self-attention and emotion-aware methods improved textual context and emotion modeling [2], [3], [4], [5].

Multimodal detection then extended the evidence domain by integrating text and visual data. Perceptual learning in modalities, hierarchical bilinear fusion, hybrid fusion, and co-attention modeling pay attention to interactions and consistency across modalities rather than handling individual modalities in isolation [1], [6], [8], [11]. Information propagation and early detection also consider aspects of information diffusion and time dynamics, and hierarchical multimodal representation gives rise to yet another method of capturing intricate relations between heterogeneous data [7], [14], [15].

The current research also proposes the use of external knowledge and graph-based structures to facilitate context-aware reasoning. Knowledge-enriched vision-language learning combines external knowledge within multimodal learning, while heterogeneous graphs leverage the connections between entities and knowledge bases [12], [13], [17]. Multilingual multimodal detection is one example of extending the detection ability to particular language-based social-media platforms, while real-time detection accounts for the computational challenges of quickly produced information [18], [19]. Systematic and meta-analytical studies also suggest that detection results are significantly affected by data set properties, modalities, and experimental design [9], [10], [16]. Table 1 therefore presents a comprehensive comparison of main research directions, technical advancements and limitations motivating future work.

Table 1. Comparison of existing methodological approaches for multimodal fake-news detection

Methodological Category	Core Technique	Major Contribution	Key Limitation
Fuzzy and Multitask Deep Learning [4], [5]	Fuzzy Deep Learning, Multitask Learning	Handles uncertainty in textual representations and jointly models detection, categorization, and risk-related tasks.	Primarily dependent on textual information and task-specific representations.

Attention and Emotion-Aware Learning [2], [3]	Self-Attention, Emotion-Aware Representation	Captures contextual dependencies and affective signals that can improve discrimination of deceptive narratives.	Limited ability to reason over visual and external contextual evidence.
Multimodal Fusion and Consistency [1], [6], [8], [11]	Modality Perception, Hierarchical Bilinear Fusion, Hybrid Fusion, Co-Attention	Integrates textual and visual representations and analyzes cross-modal interactions and consistency.	Cross-modal semantic alignment remains challenging when individually credible modalities form misleading combinations.
Propagation and Early Detection [7], [14]	Reinforced Propagation Paths, Multimodal Early Detection	Incorporates propagation patterns and complementary information to support earlier misinformation identification.	Dependent on the availability and quality of propagation or temporal information.
Hierarchical Multimodal Representation [15]	Hyperbolic Representation Learning	Models complex hierarchical relationships among multimodal representations using a non-Euclidean representation space.	Increased modeling complexity and dependence on suitable relational structure.
Knowledge-Enhanced Multimodal Learning [12]	Knowledge-Enhanced Vision-Language Modeling	Introduces external knowledge to provide additional contextual grounding for multimodal fake-news analysis.	Effectiveness depends on the relevance, completeness, and reliability of external knowledge.
Knowledge Graph and Heterogeneous GNNs [13], [17]	Knowledge Graphs, Heterogeneous Graph Learning	Exploits entity relationships and multiple knowledge sources to improve contextual and relational reasoning.	Graph construction, knowledge quality, and computational complexity remain important challenges.

Multilingual and Cost-Sensitive Detection [18]	Fact-Checked Multimodal Learning, Cost-Sensitive Classification	Extends multimodal fake-news detection to Tamil social-media content while addressing unequal classification costs.	Language-specific resources and limited cross-lingual generalization remain challenging.
Real-Time Detection [19]	Big-Data Analytics, Deep Neural Networks	Addresses high-volume misinformation streams with emphasis on real-time processing.	Maintaining detection reliability under large-scale, rapidly changing information remains difficult.
Review and Meta-Analytical Evidence [9], [10], [16]	Systematic Review, Landscape Analysis, Meta-Analysis	Consolidates datasets, modalities, AI approaches, challenges, and broader evidence across fake-news detection research.	Findings remain dependent on the heterogeneity and reporting quality of the underlying literature.

3. Datasets and Evaluation Practices

The reliability of fake news detection greatly depends on the features of the datasets employed for model training as well as the circumstances in which the detection efficiency is evaluated. The modality, language, platform, annotation technique, and context all have an impact on the behavior of the detection models [9], [10], [16]. It is especially crucial to have multimodal and multilingual datasets, as modern misinformation often involves a combination of text and images, and language specific datasets account for peculiarities of languages other than English [18]. The primary datasets and their features are described in Table 2.

3.1 Major Datasets for Fake News Detection

PHEME offers event-based Twitter conversations and veracity information for studying rumor detection and early warning systems. FakeNewsNet incorporates content information with social media information to study news properties and spread properties [9], [10]. Weibo contains content, visual, and social media information, which can be useful in studying multimodal fake news detection and interactions between multimodal content [9], [10]. Tamil Fake News dataset makes the study multimodal from a language perspective, where fact-checked social media data

and cost-sensitive learning are used [18]. The main features of these datasets are shown in Table 2.

Table 2. Characteristics of datasets used for multimodal fake-news detection

Dataset	Modality	Language	Key Characteristics	Detection Focus
PHEME [9], [10]	Text, social context	English	Event-centered Twitter conversations with veracity information	Rumor and early detection
FakeNewsNet [9], [10]	Text, social context, multimedia	English	News content integrated with social-media information	Content and propagation analysis
Weibo [9], [10]	Text, image, social context	Chinese	Multimodal social-media content with interaction information	Multimodal detection
Tamil Fake News [18]	Text, image	Tamil	Fact-checked multimodal social-media content	Cost-sensitive multimodal classification

These differences carry important methodological implications. The former type of data is suitable for analyzing temporal and propagation-oriented evidence, while the latter allows one to test the consistency between text and images. Additionally, language-dependent resources highlight the shortcomings of models trained on dominant-language datasets [9], [18]. Therefore, dataset diversity plays a significant role in the evaluation of whether a given detection strategy captures general misinformation features or learns platform, language, event, or modality-specific patterns.

3.2 Evaluation Metrics and Experimental Considerations

Fake news detection results are usually measured by accuracy, precision, recall, F1 score, and AUC-ROC. While accuracy is the overall measure of classification accuracy, precision and recall describe prediction reliability and the ability to detect fake news, respectively. The F1 score balances precision and recall, while AUC-ROC is an indicator of discrimination at various classification thresholds [9], [10], [16]. The effect of dataset differences, class distribution, modality, and experiment configuration on these metrics should be considered during comparison of results.

The year-wise trend of performance, as shown in Figure 1, helps draw a comparison of the stated prediction performance of the fake news identification task for the selected years. The trend demonstrates the overall evolution of the detection of fake news from simple content-based approaches towards advanced attention-based, multimodal, knowledge-aware, and graph-based techniques [2], [4], [6], [12], [17]. The enhanced representation learning, multimodal interactions, contextual information, and additional knowledge played a key role in enhancing the effectiveness of detecting fake news. It is important to understand that these values can be analyzed only considering the datasets used in every experiment.

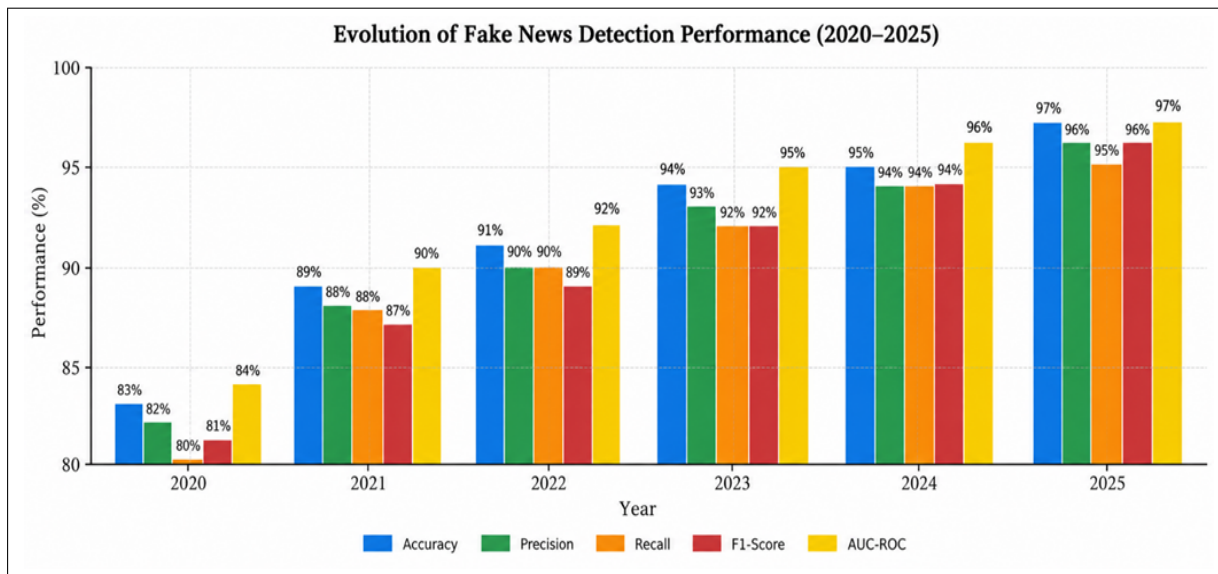


Figure 1. Year-wise fake news detection performance

4. Critical Analysis of Existing Approaches

Fake news detection methodology can be viewed as an evolution of the evidence that is employed to classify data. On one hand, text-based algorithms offer convenient representations of semantics and language; yet, the decision-making process of such systems is limited to the patterns existing within the news themselves [2], [3], [4], [5]. The multimodal approach extends the scope of the evidence by incorporating the intermodal relationships between text and visuals; therefore, such systems can detect the inconsistencies that would not be detectable using just one mode [1], [6], [8], [11].

The propagation-aware and graph-based approaches take care of another drawback by making use of the connections between the users, posts, events, and pathways through which the information is shared [7], [13], [14]. Such methods benefit from structural and

temporal information which cannot be extracted through content analysis alone. Hyperbolic embedding and knowledge-enhanced approaches provide enhanced representation of hierarchy and knowledge-based relations [12], [15]. Heterogeneous graph and LLM-inspired approaches continue along the same lines by making use of multiple knowledge bases and relational structures [13], [17]. However, their effectiveness is dependent on the graph quality and computational power.

Trade-offs can be observed throughout these types of methodologies. For text-based methodologies, there is simplicity but also little to no multimodality and context-based reasoning. Multimodal approaches give better insights but require good cross-modality alignment. Propagation and graph-based methodologies leverage the power of social relations but need enough information about those relations; the same applies to knowledge-based approaches which offer better context-based insights but rely on the completeness of the outside world [9], [10], [16].

5. Research Challenges and Open Issues

Multimodal and knowledge-aware approaches to detecting fake news have increased the amount of evidence for detecting misinformation; however, this approach has also created problems that cannot be solved by traditional approaches to classification. The key problems are related to cross-modal reasoning, dataset diversity, generalization, interpretability, knowledge reliability, and real-time scalability [1], [9], [16]. These problems are highly interrelated because the lack of diversity of datasets limits generalization, multimodal and graph structures complicate interpretability, while advanced reasoning methods make computations more intensive.

Table 3. Challenges and current research directions

Challenge	Technical Issue	Impact on Detection	Current Research Direction
Multimodal Inconsistency [1], [6]	Text and visual content may convey conflicting or misleading contextual information	Incorrect decisions when modalities appear individually credible	Cross-modal consistency and interaction modeling

Dataset Diversity [9], [18]	Limited coverage across languages, platforms, events, and modalities	Weak cross-domain and cross-language generalization	Larger, heterogeneous, and multilingual datasets
Model Interpretability [8], [16]	Complex deep and multimodal models provide limited evidence attribution	Difficult to validate and trust predictions	Explainable attention and evidence-oriented analysis
Knowledge Reliability [12], [13], [17]	External knowledge may be incomplete, outdated, or insufficiently aligned with claims	Incorrect contextual grounding	Knowledge-enhanced multimodal and graph-based reasoning
Propagation and Temporal Variation [7], [14]	Misinformation evolves rapidly and propagation evidence changes over time	Reduced effectiveness of static detection models	Early detection and propagation-aware modeling
Real-Time Scalability [19]	High-volume streams require rapid multimodal processing and inference	Increased latency and computational requirements	Efficient and scalable real-time detection

The challenges summarized in Table 3 indicate that better prediction is not enough for ensuring the reliability of misinformation detection. Future detection algorithms should also account for heterogeneity of the evidence, shifts in distributions, explainability, quality of knowledge, time dependencies, and computational aspects. Consequently, the growing complexity of the detection algorithms has to be balanced by their ability to reliably decide in the real-world environment.

6. Future Research Directions

As far as the trend for the future research in fake news detection is concerned, it is becoming more and more inclined towards the evidence-based, multimodal, explainable, and scalable fake news detection systems which do not depend on content-based classification alone. The challenges for the use of multimodal detection approaches include the development of appropriate ways of detection of semantic relations and inconsistency in the visual and textual evidence, whereas for the knowledge-augmented approaches include getting access to external

information that is valid [1], [6], [12]. Graph-based techniques can be used to facilitate reasoning about entities, users, spread patterns, and other knowledge sources, especially in the presence of context [13], [17].

The future research directions will develop robust multi-lingual and cross-domain detectors. The datasets vary greatly in terms of language, platform used, event distribution, annotation strategy, and type of modality. It will be a challenging task to generalize the results in different domains because of this variation [9], [10], [18]. Hence, future research should focus more on constructing diversified benchmarks, evaluating cross-lingual and temporally robust models, and applying standard experimental settings. Early detection needs models to make decisions based on partial observations without relying on information available after propagation.

Explanation and real-time implementation are an additional area of study. Multimodal, attention, and graph-based complex models need to give more explainable reasons for the results so that decisions can be analyzed and checked [8], [16]. On the other hand, real-time models should minimize computation costs without losing detection accuracy in a continuous flow of data [19]. Combining efficient multimodal processing, explanation reasoning, reliable grounding of knowledge, and scalable inference is likely to build a solid basis for fake news detection.

7. Conclusion

The detection of fake news has progressed from traditional text-based classification to an analytical framework, where textual, visual, contextual, propagation, and knowledge-based evidence is used. The reviewed methods provide evident evolution of fake news detection methods towards more elaborate representations of data, which would enable them to account for the complexities of modern misinformation in social media. At the same time, the differences in datasets, languages, modalities, temporal aspects, and the sources of information still influence the reliability and generalization of such methods. From this perspective, the review shows that successful fake news detection requires not only high classification accuracy but also the considerations of interpretability, evidential coherence, robustness, scalability, and the possibility of real-time implementation of methods. Thus, the evolution of such methods can be viewed as a solid basis for the development of effective and reliable approaches to fake news detection.

References

- [1] Wang, Boyue, Guangchao Wu, Xiaoyan Li, Junbin Gao, Yongli Hu, and Baocai Yin. "Modality Perception Learning-based Determinative Factor Discovery for Multimodal Fake News Detection." *IEEE Transactions on Neural Networks and Learning Systems*, 2024, vol. 36, no. 7: 12643-12654.
- [2] Jamshidi, Bahman, Saqib Hakak, and Rongxing Lu. "A Self-Attention Mechanism-based Model for Early Detection of Fake News." *IEEE Transactions on Computational Social Systems*, 2023, vol. 11, no. 4: 5241-5252.
- [3] Liu, Fei, Xinsheng Zhang, and Qi Liu. "An Emotion-Aware Approach for Fake News Detection." *IEEE Transactions on Computational Social Systems*, 2023, vol. 11, no. 3: 3516-3524.
- [4] Xu, Cheng, and M-Tahar Kechadi. "An Enhanced Fake News Detection System with Fuzzy Deep Learning." *IEEE Access*, 2024, vol. 12: 88006-88021.
- [5] Shehata, Ahmed Maher Khafaga, Mohammed Nasser Al-Suqri, Nour Eldin Mohamed Elshaiekh Osman, Faten Hamad, Yousuf Nasser Alhusaini, and Ahmed Mahfouz. "ArabFake: A Multitask Deep Learning Framework for Arabic Fake News Detection, Categorization, and risk Prediction." *IEEE Access*, 2024, vol. 12: 191345-191360.
- [6] Liu, Yahui, Wanlong Bing, Shuai Ren, and Hongliang Ma. "BC-FND: an Approach based on Hierarchical Bilinear Fusion and Multimodal Consistency for Fake news Detection." *IEEE Access*, 2024, vol. 12: 62738-62749.
- [7] Zhang, Litian, Xiaoming Zhang, Ziyi Zhou, Xi Zhang, Senzhang Wang, Philip S. Yu, and Chaozhuo Li. "Early Detection of Multimodal Fake News via Reinforced Propagation Path Generation." *IEEE Transactions on Knowledge and Data Engineering*, 2024, vol. 37, no. 2: 613-625.
- [8] Khan, Fazlullah, Ryan Alturki, Gautam Srivastava, Foziah Gazzawe, Syed Tauhid Ullah Shah, and Spyridon Mastorakis. "Explainable Detection of Fake News on Social Media using Pyramidal Co-Attention Network." *IEEE Transactions on Computational Social Systems*, 2022, vol. 11, no. 4: 4574-4583.
- [9] Hussain, Fiza Gulzar, Muhammad Wasim, Seemab Hameed, Abdur Rehman, Muhammad Nabeel Asim, and Andreas Dengel. "Fake News Detection Landscape: Datasets, Data Modalities, AI Approaches, their Challenges, and Future Perspectives." *IEEE Access*, 2025, vol. 13: 54757-54778.

- [10] Alnabhan, Mohammad Q., and Paula Branco. "Fake News Detection using Deep Learning: A Systematic Literature Review." *IEEE Access*, 2024, vol. 12: 114435-114459.
- [11] Hamed, Suhaib Kh, Mohd Juzaidin Ab Aziz, and Mohd Ridzwan Yaakub. "Improving Data Fusion for Fake News Detection: A Hybrid Fusion Approach for Unimodal and Multimodal Data." *IEEE Access*, 2024, vol. 12: 112412-112425.
- [12] Gao, Xingyu, Xi Wang, Zhenyu Chen, Wei Zhou, and Steven CH Hoi. "Knowledge Enhanced Vision and Language Model for Multi-Modal Fake News Detection." *IEEE Transactions on Multimedia*, 2024, vol. 26: 8312-8322.
- [13] Xie, Bingbing, Xiaoxiao Ma, Jia Wu, Jian Yang, and Hao Fan. "Knowledge Graph Enhanced Heterogeneous Graph Neural Network for Fake News Detection." *IEEE Transactions on Consumer Electronics*, 2023, vol. 70, no. 1: 2826-2837.
- [14] Sormeily, Asma, Sajjad Dadkhah, Xichen Zhang, and Ali A. Ghorbani. "MEFaND: A Multimodal Framework for Early Fake News Detection." *IEEE Transactions on Computational Social Systems*, 2024, vol. 11, no. 4: 5337-5353.
- [15] Feng, Shanshan, Guoxin Yu, Dawei Liu, Han Hu, Yong Luo, Hui Lin, and Yew-Soon Ong. "MHR: A Multi-Modal Hyperbolic Representation Framework for Fake News Detection." *IEEE Transactions on Knowledge and Data Engineering*, 2025, vol. 37, no. 4: 2015-2028.
- [16] Kozik, Rafał, Aleksandra Pawlicka, Marek Pawlicki, Michał Choraś, Wojciech Mazurczyk, and Krzysztof Cabaj. "A Meta-Analysis of State-of-the-Art Automated Fake News Detection Methods." *IEEE Transactions on Computational Social Systems*, 2023, vol. 11, no. 4: 5219-5229.
- [17] Xie, Bingbing, Xiaoxiao Ma, Xue Shan, Amin Beheshti, Jian Yang, Hao Fan, and Jia Wu. "Multiknowledge and LLM-Inspired Heterogeneous Graph Neural Network for Fake News Detection." *IEEE Transactions on Computational Social Systems*, 2024, vol. 12, no. 2: 682-694.
- [18] Francis, Meclin A., Ayswarya R. Kurup, B. Premjith, and Bharathi Raja Chakravarthi. "Multimodal Fake News Classification in Tamil Using Fact-Checked Social Media Content and Cost-Sensitive Learning." *IEEE Access*, 2025, vol. 13: 157477-157510
- [19] Babar, Muhammad, Awais Ahmad, Muhammad Usman Tariq, and Sarah Kaleem. "Real-Time Fake News Detection using Big Data Analytics and Deep Neural Network." *IEEE Transactions on Computational Social Systems*, 2023, vol. 11, no. 4: 5189-5198.