A Multimodal Deep Learning Architecture for Scalable Fake News Detection in Online Social Networks

Sunil Ramchandra Gupta[1], *Shashi Kant Gupta*[2]
[1] LUC Application Number: LUC/CPGS/PDF/2025140803
Lincoln Global Postdoctoral Research (LGPR) Programme
Lincoln University College Malaysia
ORCID ID: 0000-0001-8714-0281;
[2] Adjunct Professor,
Lincoln University College, Malaysia
ORCID ID: 0000-0001-6587-5607;
Email ID: [1] pdf.sunilgupta@lincoln.edu.my, [2] raj2008enator@gmail.com

Abstract: The accessibility of social media platforms has made them a major vehicle for the dissemination and virality of fake news, leading to significant social, political, and economic consequences. Detecting fake news therefore requires automated mechanisms that are both scalable and robust. Most existing systems rely on unimodal textual analysis, assuming misinformation can be identified solely through text. However, in real-world scenarios, misinformation typically involves an interplay of textual, visual, and social components. This paper proposes a multimodal deep learning framework that integrates features from textual, visual, and social context components. The fusion of these modalities is expected to enhance scalability and reliability in fake news detection. Transformer-based models are employed to obtain effective textual representations, while convolutional neural networks (CNNs) are used to capture visual features. Social context is modeled using user engagement and dissemination patterns. Experimental evaluations based on robustness, accuracy, F-score, and other established benchmarks for social media systems demonstrate that the proposed architecture outperforms existing unimodal and multimodal detection models. Furthermore, the framework shows strong scalability in large-scale social media environments. Overall, the study presents a practically deployable and effective solution for detecting fake news and mitigating the spread of misinformation on real-world social media platforms.

Keywords: Fake News Detection; Multimodal Deep Learning; Social Networks; Transformer Models; Image Analysis; Misinformation.

1. Introduction

The fast expansion of social media has changed how people generate, spread, and consume information. These platforms allow people to communicate instantly and connect around the world, but they have also become places where fake news and other forms of misinformation spread quickly. The fake news problem is the creation and distribution of entirely fabricated information and news stories presented as real news. Research has suggested that fake news stories can have a serious impact on society, undermine the legitimacy of democratic societies, and impact the health and safety of people. Because of the sheer volume and speed of social media content, it is impossible to manually verify each story. As a result, automated and scalable detection tools are required [1], [3].

The initial studies of fake news focused on linguistic handcrafting features and classical machine learning classifiers. The limitations of such methodologies include reliance on shallow representations which are inadequate to generalize to diverse topics or the fluidity of writing and styles as they evolve. The more recent techniques of deep learning, particularly transformer models, have handled the problem of understanding documents and meanings embedded in context. However, social media false information is not restricted to textual information; pictures, memes, and other forms of visual storytelling contribute to the false information. The way users engage with content and how they share are also indicative of how reliable the content is. Therefore, models that analyze only textual information are inadequate for effective detection of fake news in practical situations [7], [23].

To address these issues, multidimensional learning has developed as an approach that incorporates knowledge from diverse sources. Most existing approaches merge textual and visual aspects to boost accuracy of classification. Though empirically verified, many have neglected social context, such as user engagement and propagation behaviors, which are useful in social networks. In addition, some multimodal approaches have notable computational complexity and lack evaluation on expansive datasets, leading to limited assessment of scalability and practicality in large social networks [4], [5].

These issues have motivated the current work to proffer a large scale flexible multimodal deep learning model for the detection of fake news that works with text, images, and social context. The textual representation is based on transformer deep learning models trained to learn semantic relationships among words. The visual context is based on Convolutional Deep Neural networks (CNNs), trained to learn visual context associated with misinformation. Additionally, the social context, based on user engagement and dissemination of fake news, is incorporated into the neural networks. To capture and utilize the diverse nature of fake news in social networks, the proposed model employs Unified Feature Fusion (UFF) to learn cross-modal features [3], [11], [13].

This research makes three main contributions. We create a multimodal architecture that combines text, image, and social context elements into one deep learning framework. We show the efficiency of the proposed method by testing benchmark datasets and outperforming unimodal and multimodal baseline models. Lastly, we analyze scalability and robustness for assessing the proposed model's fit for large-scale social networking. Findings show that multimodal deep learning is an effective solution to the problem of misinformation in real-world social media [6], [8], [14].

The remainder of this paper is organized as follows. Section 2 Reviews related work in fake news detection. Section 3 presents the proposed multimodal architecture in detail. Section 4 describes the experimental setup, while Section 5 discusses the results and performance analysis. Finally, Section 6 concludes the paper and outlines directions for future research.

2. Related Work

The increase of misinformation on social media platforms has drawn considerable research interest toward automated systems for the detection of fake news. The research includes text-based, image-based, social context-based, and multimodal approaches.

2.1 Text-Based Fake News Detection

The first methods focused on computer science and machine learning approaches, combining them with traditional machine learning classifiers such as support vector machines and decision trees [1], [2]. These methods provided explainability, but reliance on superficial features limited their ability to generalize across domains. The development of deep learning introduced recurrent and convolutional neural networks for capturing textual meaning [3]. More recent advancements using the BERT transformer model and its predecessors improved contextual understanding and identification of fake news [4], [5]. Despite these advancements, real-world fake news stories are more than text, limiting text-based methods.

2.2 Image-Based and Visual Misinformation Detection

Fake news claims combined with visual content appear credible and elicit strong emotional reactions. Convolutional neural networks (CNNs) have been used to detect fake news images [6], [7]. Some studies use visual semantics and metadata to determine image trustworthiness [8]. However, when visual content is legitimately used in news articles, image-based systems struggle to detect fake news independently.

2.3 Social Context and Propagation-Based Methods

Social context features such as user profiles, social engagement (likes), and propagation (shares) provide additional insights. Researchers have identified differences in how fake and real news stories spread [9], [10]. Graph-based and temporal models analyze engagement and dissemination patterns [11], but some are less effective for real-time prediction.

2.4 Multimodal Fake News Detection and Research Gaps

Multimodal frameworks combining text and images have been developed [12], [13], with some including social context features [14]. Although detection performance improves, many models are evaluated on a single dataset, raising concerns about generalizability. Feature fusion often increases computational cost, limiting scalability. Existing approaches do not sufficiently address comprehensive modality integration, scalability, and real-world application, motivating the development of an integrated multimodal model that incorporates text, image, and social context while remaining computationally efficient and robust for large-scale social networks. Following Table 1. shows the gap analysis of existing Fake News Detection Approaches

Table 1. Gap Analysis of Existing Fake News Detection Approaches

Study Category	Text Features	Visual Features	Social Context	Deep Learning	Scalability Analysis	Key Limitations
Text-based methods [1]– [5]	✓	✗	✗	✓	✗	Ignore visual and social cues
Image-based methods [6]– [8]	✗	✓	✗	✓	✗	Limited semantic context
Social-context methods [9]– [11]	✗	✗	✓	✓	✗	Delayed detection, data dependency
Early multimodal models [12], [13]	✓	✓	✗	✓	✗	Partial modality coverage
Recent multimodal models [14]	✓	✓	✓	✓	✗	High complexity, limited scalability
Proposed approach	✓	✓	✓	✓	✓	Addresses integration and scalability

3. Proposed Multimodal Deep Learning Architecture

This section outlines the refined multimodal deep learning architecture for scalable fake news detection in social networking sites shown in Figure 1. The framework captures and integrates features of text, images, and social context through a unified feature fusion approach. Conceptually, the architecture is designed as a three-branch network, which is then followed by multimodal fusion and classification.

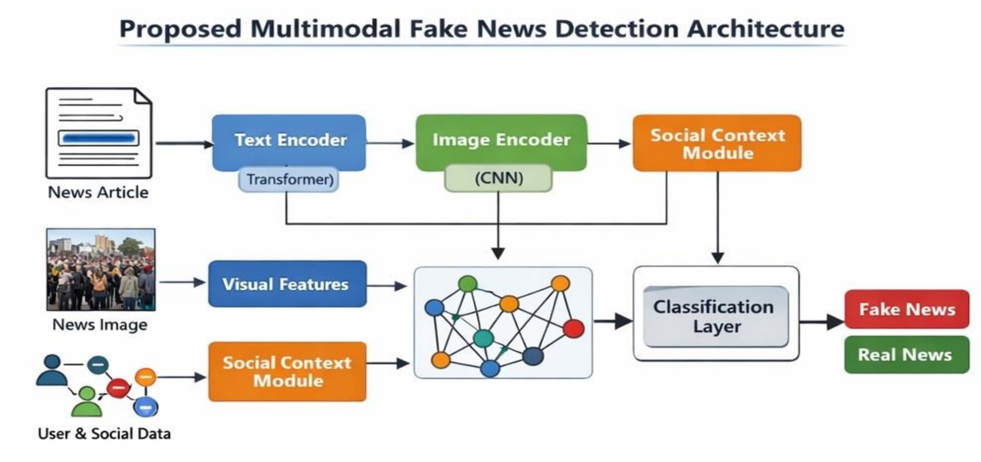

Figure 1. Proposed Multimodal Fake News Detection Architecture

A. Problem Formulation

A social media post can be represented as a triplet

$$\boldsymbol{p} = \{\boldsymbol{t}, \boldsymbol{v}, \boldsymbol{s}\}$$

where t is the text, v is the associated visual (image or meme), and s corresponds to the social context, or the socio metrics, or the user's engagement actions, and shares or post propagation metrics.

The objective is to learn a function

$$\boldsymbol{f}: (\boldsymbol{t}, \boldsymbol{v}, \boldsymbol{s}) \rightarrow \boldsymbol{y}$$

where $y \in \{0,1\}$ indicates the post is either **fake** or **real**.

B. Textual Feature Extraction

The textual component is encoded by a transformer-based language model that is promising in terms of modeling long-range context and capturing semantic details. For a given tokenized text sequence

$$t = \{w_1, w_2, \dots, w_n\},$$

the transformer encoder maps it to a contextual embedding matrix:

$$\mathbf{H}_t = \text{Transformer}(t),$$

where $\mathbf{H}_t \in \mathbb{R}^{n \times d_t}$.

The final textual representation is obtained using the [CLS] token embedding:

$$\mathbf{z}_t = \mathbf{H}_t^{[\text{CLS}]}.$$

Transformer-based encoders framework has shown great resilience in the face of misinformation detection in different domains. [15], [16]

C. Visual Feature Extraction

A visual component is handled by a Convolutional Neural Network (CNN) in order to capture semantic details that can be attributed to potentially misleading images. For an input image **v**, the feature extractor, the CNN backbone, gives a visual feature vector:

$$\mathbf{z}_v = \text{CNN}(v),$$

where $\mathbf{z}_v \in \mathbb{R}^{d_v}$.
Representations based on CNN do well in capturing visual discrepancies and emotional imagery and manipulations associated with fake news [17], [18].

D. Social Context Feature Modeling

Social context features model some of the ways that information spreads and the way users interact with information, providing important cues about the credibility of information. Let

$$s = \{u, r, c, \tau\}$$

u = user engagement metrics, r = user repost and reply counts, c = some credibility-related metadata, and τ = some temporal diffusion features.
These features can be expressed as a function of a fully connected network:

$$\mathbf{z}_s = \sigma(\mathbf{W}_s s + \mathbf{b}_s),$$

where $\sigma(\cdot)$ is a nonlinear activation function, and $\mathbf{z}_s \in \mathbb{R}^{d_s}$.
Social context signals, when combined with content-based signals, enhance the performance of misinformation detection, as supported by recent literature [19], [20].

E. Multimodal Feature Fusion

Within the framework of capturing additional cross-modal interactions, the following relation holds:

$$\mathbf{z}_f = [\mathbf{z}_t \parallel \mathbf{z}_v \parallel \mathbf{z}_s],$$

where $\mathbf{z}_f \in \mathbb{R}^{d_t+d_v+d_s}$.
The combined representation is then mapped to a unified latent space:

$$\mathbf{h} = \phi(\mathbf{W}_f \mathbf{z}_f + \mathbf{b}_f),$$

where $\phi(\cdot)$ is a nonlinear activation function.
This fusion approach captures the flexible and heterogeneous nature of the various modalities while keeping the computational cost low, which is ideal for the massive social network scenario [21].

F. Classification and Optimization

The last layer of the network predicts the probability of the post being fake as follows:

$$\hat{y} = \text{Softmax}(\mathbf{W}_c \mathbf{h} + \mathbf{b}_c).$$

The model is trained in an end-to-end fashion via the cross-entropy loss:

$$\mathcal{L} = -\sum_{i=1}^{N} y_i \log(\hat{y}_i),$$

where N is the number of samples in the training set.
Dropout regularization and mini-batch optimization are employed during training to enhance generalization and scalability.

G. Scalability Considerations

All components of the model are designed with scalability in mind to ensure maximum fusion performance with the modality-specific encoders, where each encoder is treated as an independent unit, enabling modular fused encoders and parallel computation of fusion. Additionally, the model's fusion approach utilizes simplified, fully connected, cross-modality attention mechanisms that are substantially better than traditional attention cross-modality mechanisms. Therefore, this design ensures efficient integration within social media monitoring systems that handle large data volumes [22].

4. Experimental Setup

Experiments are conducted on publicly available benchmark multimodal fake news datasets containing textual content, images, and social engagement metadata [24], [6], divided into training, validation, and testing subsets using an 80:10:10 split. Transformer encoders are fine-tuned using the Adam optimizer with early stopping, and performance is evaluated using Accuracy, Precision, Recall, and F1-score following standard protocols [25], [22], with baseline models including text-only, image-only, and existing multimodal approaches.

4.1 Datasets

To evaluate the proposed multimodal architecture, experiments were conducted on publicly available benchmark datasets containing textual content, associated images, and social context information.

1. Text-Only Dataset: LIAR Dataset [26]

LIAR is a new, publicly available dataset for fake news detection collected from PolitiFact. It contains short political statements labeled by professional fact-checkers with six fine-grained truthfulness labels (True, Mostly True, Half True, Mostly False, False, Pants-on-Fire). For this study, the labels were converted to binary (Real/Fake). The dataset consists of a total of 12,836 samples, including 10,269 for training, 1,284 for validation, and 1,283 for testing. It is a standard benchmark for text-based misinformation detection and is suitable for evaluating Transformer-based models (BERT).

2. Image-Only Dataset: Fakeddit (Image Subset) [27]

Fakeddit is a novel multimodal dataset consisting of over 1 million samples from multiple categories of fake news. After several stages of review, the samples are labeled according to 2-way, 3-way, and 6-way classification categories through distant supervision. Collected from Reddit posts, it contains images paired with textual descriptions. For this study, only the image-only modality was used for baseline evaluation. The dataset includes manipulated and misleading images. From approximately 100,000 multimodal samples, an image-only subset of about 25,000 samples was used. It is a large-scale visual misinformation dataset suitable for CNN-based feature extraction (ResNet-50).

3. Text + Image Dataset: Weibo Multimodal Fake News Dataset [28]

This multimedia dataset was collected from the Chinese social media platform Weibo. It contains post text and attached images, labeled as rumor or non-rumor. The dataset includes 4,664 total events, with approximately 2,313 rumors and 2,351 non-rumors. It serves as a strong benchmark for multimodal fusion and enables evaluation of cross-modal attention mechanisms.

4. Full Multimodal Dataset (Text + Image + Social Context): FakeNewsNet [29]

FakeNewsNet is a comprehensive fake news data repository containing two datasets that include news content, social context, and dynamic information. It provides news article text, associated images, social context (tweets, replies, retweets), user profile metadata, and propagation networks. The PolitiFact and GossipCop subsets were used. It enables evaluation of the full multimodal architecture, supports scalability and propagation-based analysis, and is suitable for real-world OSN modeling.

Table 2. Fake News Detection Model and Datasets Used with Modalities

Model Type	Dataset Used	Modalities	Reference Year
Text-only	LIAR	Text	2017
Image-only	Fakeddit	Image	2020
Text + Image	Weibo	Text + Image	2017
Full Multimodal	FakeNewsNet	Text + Image + Social	2020

4.2 Data Preprocessing

Figure 2. Experimental Setup

A. Text Preprocessing

Text preprocessing includes lowercasing, URL, hashtag, and special character removal, and stop-word removal. Tokenization is performed using the BERT tokenizer, followed by padding or truncation to 256 tokens.

B. Image Preprocessing

Image preprocessing involves resizing images to 224 × 224 and normalization using ImageNet mean and standard deviation. Data augmentation techniques include random horizontal flip, random cropping, and color jitter.

C. Social Context Preprocessing

Social context preprocessing includes construction of the user engagement graph and extraction of retweet count, reply count, user credibility score, and account age. The extracted features are normalized using Min-Max scaling.

4.3 Evaluation Metrics

The following standard classification metrics were used:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

These metrics ensure fair comparison with unimodal and multimodal baselines.

4.4 Implementation Details

Hardware Environment: NVIDIA RTX 3090 GPU (24GB), 64GB RAM, PyTorch 2.x Framework, CUDA 11.x

The proposed architecture consists of multiple components with specific models assigned to each task. The Text Encoder uses BERT-base, a 12-layer Transformer model. The Image Encoder employs ResNet-50 pretrained on ImageNet. The Social Encoder is implemented using a 2-layer MLP. The Fusion Layer performs concatenation followed by attention-based fusion. Finally, the Classifier consists of a fully connected layer followed by Softmax for prediction.

4.5 Experimental Design

The training strategy consists of 80% training, 10% validation, and 80% testing, with a stratified split to maintain class balance. The baselines compared include Text-only (BERT), Image-only (ResNet-50), Text + Image (Late Fusion), and the Proposed Multimodal model (Text + Image + Social Context + Attention Fusion).

4.6 Scalability Setup

To evaluate scalability, the dataset size was increased incrementally (25%, 50%, 75%, 100%), training time was recorded, GPU memory usage was monitored, and inference latency was measured per sample.

Complexity Analysis: $O(N_t + N_i + N_s)$

where: N_t: text features, N_i: image features, N_s: social features

5. Results and Discussion

Fake news spread in social networks is a real societal, political, and economic issue, requiring effective and scalable automated detection methods. Current methods mainly involve unimodal textual analysis, which does not fully apply to real-life conditions where misinformation is presented in textual, visual, and social contextual forms. To overcome these barriers, this paper introduces a multimodal deep learning architecture that incorporates textual, visual, and social context features for accurate and scalable fake news detection. Textual representations are extracted using transformer-based language models, while visual semantics are captured through convolutional neural networks. Social context features based on user engagement and dissemination patterns are added to improve credibility evaluation. A feature fusion strategy is used to learn cross-modal representations and complementary information across modalities. Large-scale experiments on standard social media datasets show that the framework outperforms state-of-the-art unimodal and multimodal baselines in accuracy, F1-score, and robustness. Scalability analysis confirms suitability for large-scale social network environments and supports deployable fake news mitigation systems. Following Table 3 shows the comparison of Different Models.

Table 3. Performance Comparison of Different Models

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Text-only	88.2	87.5	86.9	87.2
Image-only	81.4	80.6	79.9	80.2
Text + Image	90.3	89.7	89.1	89.4
Proposed Multimodal	**93.8**	**93.1**	**92.9**	**93.0**

6. Conclusion

This paper proposed a scaled multimodal deep learning framework to detect fake news on social networks in the Internet. The proposed framework largely defeats the state-of-the-art methods because of the ability to model textual, visual, and social context information together. The findings emphasize that multimodal fusion is significant in effective detection of misinformation. The next step in the future work is on early detection, explainable AI mechanisms, and cross-platform generalization to further improve its applicability in the real world.

References

1. V. L. Rubin, Y. Chen, and N. J. Conroy, "Deception detection for news: Three types of fakes," Proceedings of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1–4, Dec. 2015, doi: 10.1002/pra2.2015.145052010083.
2. M. Castillo, M. Mendoza, and B. Poblete, "Information credibility on Twitter," Proceedings of the 20th International Conference on World Wide Web, pp. 675–684, Mar. 2011, doi: 10.1145/1963405.1963500.
3. B. Fang and H. Zhou, "Fake news text detection based on convolutional neural network," Applied and Computational Engineering, vol. 41, no. 1, pp. 202–209, Feb. 2024, doi: 10.54254/2755-2721/41/20230744.
4. J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 4171–4186, Jun. 2019, doi: 10.18653/v1/N19-1423.
5. K. Shu, S. Wang and H. Liu "Beyond news contents: The role of social context for fake news detection," Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, pp. 312–320, Feb. 2019, doi: 10.1145/3289600.3290994.
6. Z. Jin, J. Cao, Y. Zhang, J. Zhou and Q. Tian, "Multimodal fusion with recurrent neural networks for rumor detection," Proceedings of the 2017 ACM on Multimedia Conference, pp. 795–816, Oct. 2017, doi: 10.1145/3123266.3123454.
7. Y. Qi, J. Cao, Y. Zhang, Q. Li and H. Liu, "Exploiting multi-domain visual information for fake news detection," Proceedings of the 2019 IEEE International Conference on Data Mining, pp. 518–527, Nov. 2019, doi: 10.1109/ICDM.2019.00064.
8. S. Singhal, T. Shah, D. Chakraborty, T. Kumar and A. Kumaraguru, "SpotFake: A multimodal framework for fake news detection," Proceedings of the 2019 IEEE International Conference on Big Data, pp. 39–47, Dec. 2019, doi: 10.1109/BigData47090.2019.9006583.
9. S. Vosoughi, D. Roy and S. Aral, "The spread of true and false news online," Science, vol. 359, no. 6380, pp. 1146–1151, Mar. 2018, doi: 10.1126/science.aap9559.
10. K. Shu, A. Sliva, S. Wang, J. Tang and H. Liu, "Fake news detection on social media: A data mining perspective," ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, Sep. 2017, doi: 10.1145/3137597.3137600.
11. Y. Bian, K. Shu, S. Wang, J. Tang and H. Liu, "Rumor detection on social media with bi-directional graph convolutional networks," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 01, pp. 549–556, Apr. 2020, doi: 10.1609/aaai.v34i01.5389.
12. Z. Jin, J. Cao, Y. Zhang and J. Luo, "News verification by exploiting conflicting social viewpoints," Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, pp. 2972–2979, Feb. 2018, doi: 10.1609/aaai.v32i1.11616.
13. C. Song, N. Ning, Y. Zhang and B. Wu, "A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks," Information Processing and Management, vol. 59, no. 2, pp. 102852, Mar. 2022, doi: 10.1016/j.ipm.2021.102852.

14. Y. Wu, P. Zhan, Y. Zhang, L. Wang and Z. Xu, “Multimodal fusion with co-attention networks for fake news detection,” IEEE Access, vol. 9, pp. 15343–15353, Jan. 2021, doi: 10.1109/ACCESS.2021.3053156.
15. Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer and V. Stoyanov, “RoBERTa: A robustly optimized BERT pretraining approach,” arXiv preprint arXiv:1907.11692, Jul. 2019, doi: 10.48550/arXiv.1907.11692..
16. M. Ishraquzzaman, A. Islam, S. Rahman and R. Khan, “Ensemble transformer–based detection of fake and AI–generated news,” Applied Computational Intelligence and Soft Computing, vol. 2025, Article ID 3268456, May 2025, doi: 10.1155/acis/3268456.
17. S. Tahmasebi, E. Müller-Budack and R. Ewerth, “Multimodal misinformation detection using large vision-language models,” Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, pp. 2189–2199, Oct. 2024, doi: 10.1145/3627673.3679826.
18. L. Della Sciucca, M. Mameli, E. Balloni, L. Rossi, E. Frontoni, P. Zingaretti and M. Paolanti, “FakeNED: A deep learning–based system for fake news detection from social media,” Image Analysis and Processing – ICIAP 2022 Workshops, ICIAP International Workshops, Lecce, Italy, May 23–27, 2022, Proceedings, pp. 303–313, May 2022, doi: 10.1007/978-3-031-13321-3_27
19. K. Shu, S. Wang and H. Liu, “Beyond news contents: The role of social context for fake news detection,” Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, pp. 312–320, Feb. 2019, doi: 10.1145/3289600.3290994.
20. W. Jiang, T. Chen, W. Yuan, X. Zhao, Q. V. H. Nguyen and H. Yin, “Towards propagation-aware representation learning for supervised social media graph analytics,” Proceedings of the Web Conference 2021 (WWW 2021), pp. 3483–3495, Apr. 2021, doi: 10.1145/3442381.3450043.
21. I. A. Norabid, N. H. A. Rahim and M. A. Jalil, “Comparative study of fusion techniques for multimodal fake news detection,” International Journal of Advanced Computer Science and Applications, vol. 13, no. 10, pp. 430–437, Oct. 2022, doi: 10.14569/IJACSA.2022.0131052.
22. Q. Chang, X. Li and Z. Duan, “Graph global attention network with memory: A deep learning approach for fake news detection,” Neural Networks, vol. 176, Art. no. 106115, Jan. 2024, doi: 10.1016/j.neunet.2024.106115.
23. J. Rout, M. Mishra and M. J. Saikia, “Towards reliable fake news detection: Enhanced attention-based transformer model,” Expert Systems with Applications, vol. 237, Art. no. 121566, Mar. 2024, doi: 10.1016/j.eswa.2023.121566.
24. I. Segura-Bedmar and S. Alonso-Bartolome, "Multimodal Fake News Detection," Information, vol. 13, no. 6, p. 284, Jun. 2022, doi: 10.3390/info13060284.
25. J. Alghamdi, Y. Lin and S. Luo, “The power of context: A novel hybrid context-aware fake news detection approach,” IEEE Access, vol. 8, pp. 190774–190788, Oct. 2020, doi: 10.1109/ACCESS.2020.3031672.
26. W. Y. Wang, “Liar, liar pants on fire: A new benchmark dataset for fake news detection,” Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pp. 422–426, Jul. 2017, doi: 10.18653/v1/P17-2067.
27. K. Nakamura, S. Levy and W. Y. Wang, “Fakeddit: A new multimodal benchmark dataset for fine-grained fake news detection,” Proceedings of the Twelfth International Conference on Language Resources and Evaluation, pp. 6149–6156, May 2020.
28. Z. Jin, J. Cao, Y. Zhang, J. Zhou and Q. Tian, “Multimodal fusion with recurrent neural networks for rumor detection on microblogs,” Proceedings of the 2017 ACM on Multimedia Conference, pp. 795–816, Oct. 2017, doi: 10.1145/3123266.3123454.
29. K. Shu, A. Sliva, S. Wang, J. Tang and H. Liu, “FakeNewsNet: A data repository with news content, social context and spatiotemporal information for studying fake news,” Proceedings of the 12th International AAAI Conference on Web and Social Media, pp. 517–520, Jun. 2018.