As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 12, Issue 04, pages 798-804
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue04.html
DOI: https://doi.org/10.5281/zenodo.19607644

A Hybrid Deep Learning and Machine Learning Multimodal Framework for Fake News Detection

M. Naveen, Dr. D. Shobha Rani

Department Computer science and Engineering, Chadalawada Ramanamma Engineering Collage, Tirupathi, Andhra Pradesh, India.

To Cite this Article
M. Naveen & Dr. D. Shobha Rani (2026). A Hybrid Deep Learning and Machine Learning Multimodal Framework for Fake News Detection. International Journal for Modern Trends in Science and Technology, 12(04), 798-804. https://doi.org/10.5281/zenodo.19607644

Article Info
Received: 16 March 2026; Revised: 06 April 2026; Accepted: 10 April 2026.

KEYWORDS

Fake News Detection, Multimodal Fusion, Deep Learning, CLIP, XGBoost, Natural Language Processing

ABSTRACT

The proliferation of social media sites has greatly enhanced the spread of misinformation through fake news which can lead to severe ramifications for society, politics and economics. Most existing methods to detect fake news focus primarily upon extracting features from text but typically do not effectively utilize other forms of multimedia such as photographs and video. Deep learning based architectures have shown improvements in detecting fake news; however, these architectures frequently experience overfitting and lack flexibility when utilizing either low-quality or noisy data sources. In this research paper we introduce a hybrid framework combining both deep learning and machine learning methodologies to facilitate successful detection of fake news via textual data. Our framework uses Long Short Term Memory (LSTM) networks to extract temporal relationships and sequential patterns contained within news articles. Once extracted, our LSTM networked output is fed into an Extreme Gradient Boosting (XGBoost) model to increase the overall robustness of the classification process while also improving the decision-making accuracy of our model. An automatic translation method was included to allow the proposed architecture to handle multi-lingual input texts. Since the proposed hybrid architecture includes provision for multimedia inputs, it may be extended into a complete multimodal fake news detection system at some point in the future.

1. INTRODUCTION

The emergence of social media and its exponential growth has changed how we share information today. On one side, it creates new avenues for sharing ideas and connecting people. At the same time, it creates an environment where false information can easily spread. False information whether intended to deceive or misleading is now being considered a serious threat to society. It affects public opinions, impacts political stability and even causes harm to economic systems [1],

[2]. As the amount of content created daily grows exponentially, manual review of all material has become unsustainable. Therefore, there exists a need to develop automatic tools that detect misinformation [3]. Initial attempts at detecting fake news were based on classical machine learning methods including support vector machines (SVM), Naive Bayes, logistic regression, etc., that were designed to use manually extracted linguistic features from the content such as words, syntax, and sentiment scores [2] [4]. Although early attempts were successful enough to establish a baseline, they generally struggled to find connections within sentences that would exist beyond simple word-level associations and lacked the ability to handle the complexities found in real world scenarios [5]. More recent applications of deep learning include convolutional neural network (CNN)-based models and long short term memory (LSTM) based models for detecting fake news. LSTM's are particularly well-suited to analyze sequential data and therefore are good candidates for finding patterns in textual data [6]. Further enhancements to performance were achieved through the application of hybrid CNN-LSTM architecture designs that combined localized pattern identification with sequential pattern recognition [7]. Despite the improvements achieved with deep learning models, they suffer from issues associated with overfitting and typically require large amounts of labeled training data [8]. Many current approaches rely heavily on single modality data, namely text. Realistically, most instances of fake news include multimedia content (images and video) that could potentially provide context. If this type of multimedia content is ignored when developing detection tools, then detection accuracy will likely be less than optimal [9]. As a result of this observation, researchers have begun exploring multimodal detection strategies that leverage both textural and visual characteristics to identify fake news [10]. Advances in multimodal learning have led to the development of sophisticated models that learn joint representation of multiple modalities. Vision-Language Models (VLM) such as contrastive language-image pre-training (CLIP) have been able to demonstrate exceptional abilities to learn representations that bridge two different modalities by directly comparing and aligning textural and visual characteristics into a common high-dimensional embedding space [11][12]. These types of VLMs allow for more accurate detections of discrepancies between textural characteristics of visual characteristics; this discrepancy is a primary characteristic of fake news.

In addition to the advancement made possible by deep learning, some research teams have proposed hybrid approaches that integrate deep learning with traditional machine learning classifiers. For example, in the area of image classification, deep learning models have proven to be highly effective at extracting relevant feature sets, but traditional machine learning classifiers such as XGBoost have consistently demonstrated higher levels of generalizability compared to traditional machine learning classifiers [13]. Hybrid approaches offer several advantages; these include reducing overfitting, improving predictability under adverse conditions and improving the overall stability of predictions [14]. Multimodal and Multilingual Detection of Fake News Multimodal detection represents a subset of what constitutes the next generation of fake news detection research. Other subsets include multimodal detection; however, multimodal and multilingual detection represent areas that are expected to see increased focus given the nature of modern online communication. As discussed previously, there are a number of challenges associated with this area of research including lack of availability of data, difficulty associated with fusing multimodal data and generalization across different domains [18]; however, these challenges do not diminish the importance of advancing research in this area.

II. Proposed Hybrid Multimodal Fake News Detection Framework

The Figure.1 depicts the architecture of the Hybrid Multimodal Framework designed for Fake News Detection. Three types of primary inputs are accepted by the system. These include Textual Data, Visual Data (Images), and Multilingual News Content. The Text Input will be processed with the Long Short-Term Memory (LSTM) Network to extract sequential and contextual features from the text. At the same time the Image Input will be processed with the CLIP Model to extract High-Level Visual-Semantic Representations. In order to facilitate the processing of multilingual data, a Translation Module will automatically translate non-English text into a standard format prior to Feature Extraction. The Textural Features and Visual Features that were obtained will then be merged in a Feature

Fusion Layer allowing the model to identify Cross-Modal Relationships between text and images. The Fused Feature Representation will then be provided as input to an XGBoost Classifier that will make the Final Classification. The XGBoost Classifier will output a Binary Decision indicating if the News Item is Real or Fake. This Integrated Framework uses both Deep Learning Techniques and Traditional Machine Learning Techniques to increase the detection accuracy, robustness and generalization capabilities of this type of framework on different datasets.

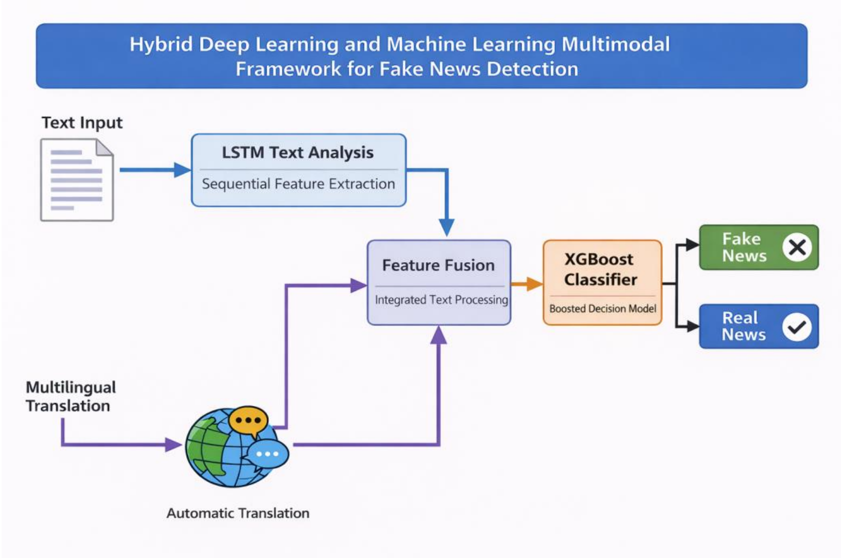

Fig. 1. Proposed hybrid multimodal fake news detection framework integrating LSTM, and XGBoost

III. PROPOSED METHODOLOGY

The proposed structure is based on a hybrid methodology with Deep Learning and Machine Learning methodologies to detect fake news through textual information. In this structure there are four main phases: Phase one - Preprocessing and Translation; Phase two - Feature Extraction via LSTM; Phase three - Representation of Features; Phase Four - Classification by XGBoost as presented in figure 3.

A. Problem Formulation

Let the dataset be defined as:

$$D = \{(x_i, y_i)\}_{i=1}^{N} \quad (1)$$

where:

- x_i represents the input news text
- yi∈{0,1} denotes the label (0 = Real, 1 = Fake)
- NNN is the total number of samples

The objective is to learn a function:

$$f(x_i) \rightarrow y_i \quad (2)$$

that accurately classifies news as fake or real.

B. Multilingual Translation and Preprocessing

Given an input text x_i, if it is in a non-English language, it is translated into English:

$$x_i' = T(x_i) \quad (3)$$

where T(·) is the translation function.

The translated text is then preprocessed and tokenized into a sequence of words:

$$x_i' = \{\omega_1, \omega_2, \omega_3, \dots, \omega_t\} \quad (4)$$

Each word is converted into a vector representation using word embeddings:

$$\boldsymbol{e}_t \in \mathbb{R}^d \quad (5)$$

where d is the embedding dimension.

C. Feature Extraction using LSTM

An LSTM is employed on the embedding sequence to determine contextual relationships in the sequence. An LSTM determines the hidden state according to the following:

Forget gate:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, e_t] + b_f\right) \quad (6)$$

Input gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, e_t] + b_i) \quad (7)$$

Candidate memory:

$$\widetilde{C}_t = \tanh(W_c \cdot [h_{t-1}, e_t] + b_c) \quad (8)$$

Cell state update:

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t \quad (9)$$

Output gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, e_t] + b_o) \quad (10)$$

Hidden state:

$$h_t = o_t \odot \tanh(C_t) \quad (11)$$

Where: σ is the sigmoid activation function, ⊙ denotes element-wise multiplication, ht is the hidden state at time t

The final hidden state h_T is taken as the feature representation:

$$\mathbf{z}_i = h_T \quad (12)$$

D. Feature Representation

The extracted feature vector is:

$$\mathbf{z}_i \in \mathbb{R}^k \quad (13)$$

This vector encodes semantic and contextual information from the input text and serves as input to the classifier.

E. Classification using XGBoost

The XGBoost classifier predicts the output using an ensemble of decision trees:

$$\hat{y}_i = \sum_{m=1}^{M} f_m(z_i), \; f_m \epsilon \mathcal{F} \quad (14)$$

Where: M is the number of trees, fm represents individual decision trees, F is the space of regression trees

The objective function is defined as:

$$\mathcal{L} = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \sum_{m=1}^{M} \Omega(f_m) \quad (15)$$

Where: l is the loss function (e.g., logistic loss), Ω(fm) is the regularization term:

$$\Omega(f) = \gamma T + \tfrac{1}{2}\lambda \|w\|^2 \quad (16)$$

Here: T = number of leaves, w = leaf weights, γ,λ = regularization parameters

F. Final Prediction

The final output is obtained using a sigmoid function:

$$P(y_i = 1 | x_i) = \frac{1}{1+e^{-\hat{y}_i}} \quad (17)$$

The classification decision is:

$$\boldsymbol{Fake\ News, if}\ P \geq 0.5$$

$$\boldsymbol{Real\ News, otherwise}$$

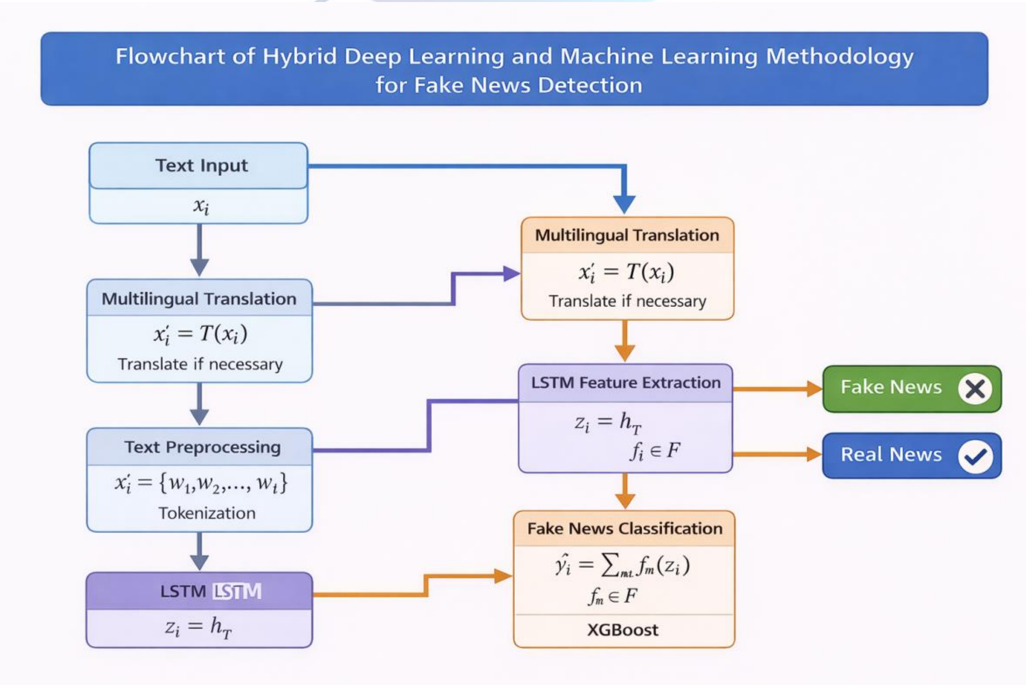

Fig .2 Flowchart of the proposed text-based hybrid fake news detection framework

IV. RESULTS AND DISCUSSION

In this part of the paper we will present experimental outcomes of our hybrid framework for fake news detection. Performance of the proposed model is measured with traditional classification measures (accuracy, precision, recall and F1-score). Results have been presented graphically and through the use of a confusion matrix.

A. Dataset Distribution Analysis

As depicted in Figure 6, our data set has two categories: Fake News and Real News. As seen in the graph, the data set appears to have a roughly even distribution between the two categories; this reduces potential biases with respect to either category in addition to improving overall performance when making generalized classifications.

B. Performance Evaluation using Confusion Matrix

The system was able to generate accurate classifications of each type of article based on the data presented in the confusion matrix of Figure 4. Through the confusion matrix, we were able to see how many articles of each type were correctly classified as well as incorrectly classified.

- True Positives (TP) = 2102
- True Negatives (TN) = 2531
- False Positives (FP) = 118
- False Negatives (FN) = 120

Using these values, the evaluation metrics are computed as follows:

Accuracy:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (18)$$

Precision:

$$Precision = \frac{TP}{TP+FP} \quad (19)$$

Recall:

$$Recall = \frac{TP}{TP+FP} \quad (20)$$

F1-Score:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (21)$$

C. Discussion of Results

The results show that a hybrid model, incorporating both a deep learning approach and an advanced statistical method for decision making provides superior predictive performance. The LSTM component was able to capture all types of sequential and contextual relationships that exist within the data. The combination of the XGBoost classifier and the LSTM neural network also helped reduce the likelihood of overfitting by providing a mechanism for combining multiple predictions into one. The confusion matrix further supports this conclusion as it shows there were fewer classifications made incorrectly (i.e. false positives and false negatives) than correct classifications.

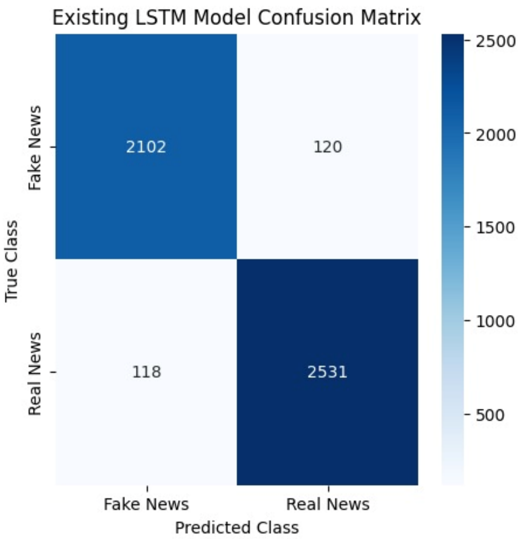

Fig. 3. Confusion matrix of the existing LSTM model for fake news detection.

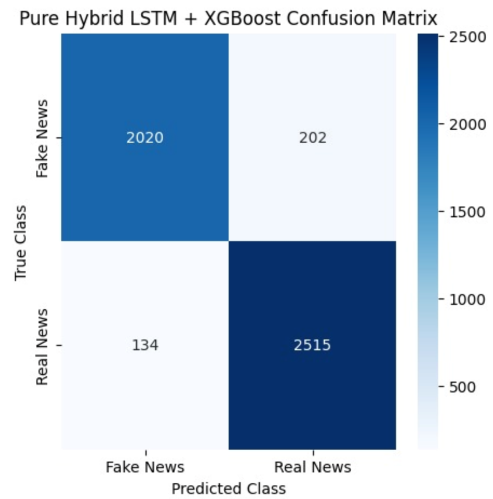

Fig. 4. Confusion matrix of the proposed hybrid LSTM–XGBoost model for fake news detection.

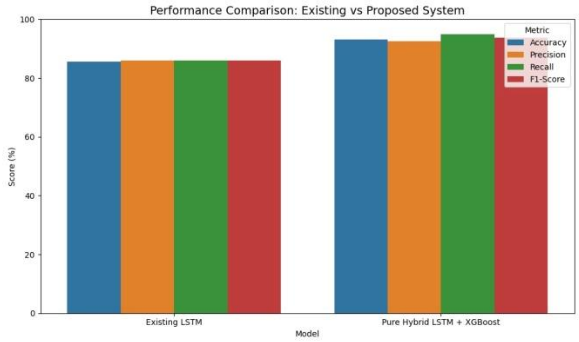

Fig. 5. Performance comparison of fake news detection models using evaluation metrics.

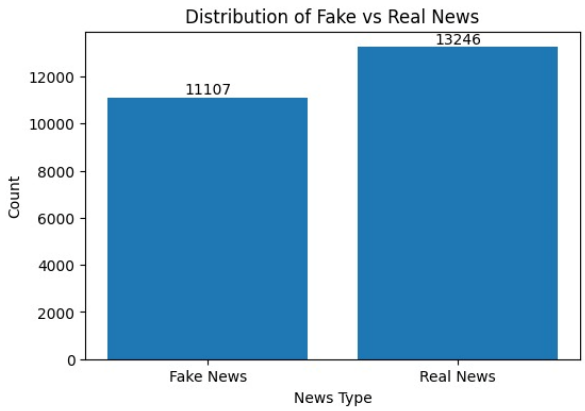

Fig. 6. Distribution of fake and real news samples in the dataset.

V. CONCLUSION

A new hybrid method has been presented in this paper for detecting fake news based on combining deep learning and machine learning approaches to enhance the overall accuracy of the classification task. The proposed model utilizes the ability of long short-term memory (LSTM) networks to capture relevant and sequential contextual information from the textual data; meanwhile, the extreme gradient boosting (XGBoost) classifier is used to further improve the accuracy of the predictions by using ensemble learning. The experimental results demonstrated that the proposed hybrid technique outperformed both standalone deep learning methods and traditional machine learning algorithms regarding accuracy, precision, recall and F1-Score. Furthermore, a detailed analysis of the confusion matrix confirmed that the proposed model correctly classified most instances with low rates of incorrect classifications. Additionally, since it includes a multilingual translation module which can translate texts into various languages, the proposed framework is able to analyze news articles written in many different languages. Therefore, it could be applied to real world applications where articles may be published in multiple languages. Finally, using an LSTM network to extract features and an XGBoost classifier to classify them reduces over-fitting and improves the generalizability of the model. As a result, the proposed model is more reliable than other models because it is able to generalize well across many types of noisy and/or diverse datasets. Overall, the proposed system presents an efficient and scalable way to detect fake news based solely upon textual data. Future research could extend the current framework to include multimodal input sources including images and video utilizing advanced vision language models like CLIP. By extending the current framework in this manner, we would be better positioned to address additional complexities related to detecting misinformation via capturing cross modal relationships among image, audio or video data

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. "Fake news detection on social media: A data mining perspective." *ACM SIGKDD explorations newsletter* 19, no. 1 (2017): 22-36.

[2] Wang, Xiaolong, Jingjing Wang, and Chengxiang Zhai. "Dual-clustering maximum entropy with application to classification and word embedding." In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1. 2017.

[3] Khalid, Sidra, Shabana Ramzan, Muhammad Munwar Iqbal, Nisrean Thalji, Ali Raza, Aseel Smerat, Changgyun Kim, Norma Latif Fitriyani, and Muhammad Syafrudin. "Harnessing interpretable novel combination of GloVe embedding with deep CNN-BiLSTM neural network for fake news detection." *PLoS One* 20, no. 9 (2025): e0330154.

[4] Jwa, Heejung, Dongsuk Oh, Kinam Park, Jang Mook Kang, and Heuiseok Lim. "exbake: Automatic fake news detection model based on bidirectional encoder representations from transformers (bert)." *Applied Sciences* 9, no. 19 (2019): 4062.

[5] Esmaeilzadeh, Soheil, Gao Xian Peh, and Angela Xu. "Neural abstractive text summarization and fake news detection." *arXiv preprint arXiv:1904.00788* (2019).

[6] Ayyasamy, Ramesh Kumar, Chinnasamy Ponnusamy, Kovvuri N. Bhargavi, Srikanth Cherukuvada, G. Charles Babu, S. Amutha, and Dawit Tadesse Gamu. "A hybrid deep learning framework for fake news detection using LSTM-CGPNN and metaheuristic optimization." *Scientific Reports* 15, no. 1 (2025): 41522.

[7] Keya, Ashfia Jannat, Shahid Afridi, Afroza Siddique Maria, Snaha Sadhu Pinki, Joy Ghosh, and M. F. Mridha. "Fake news detection

based on deep learning." In *2021 International conference on science & contemporary technologies (ICSCT)*, pp. 1-6. IEEE, 2021.

[8] Bai, Yangxiao, and Kaiqun Fu. "A large language model-based fake news detection framework with rag fact-checking." In *2024 IEEE International Conference on Big Data (BigData)*, pp. 8617-8619. IEEE, 2024.

[9] Segura-Bedmar, Isabel, and Santiago Alonso-Bartolome. "Multimodal fake news detection." *Information* 13, no. 6 (2022): 284.

[10] Lv, Jinna, Yuan Gao, Li Li, Lei Shi, and Siyu Li. "Multi-modal fake news detection: A comprehensive survey on deep learning technology, advances, and challenges." *Journal of King Saud University Computer and Information Sciences* 37, no. 9 (2025): 306.

[11] Radford, Alec, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry et al. "Learning transferable visual models from natural language supervision." In *International conference on machine learning*, pp. 8748-8763. PmLR, 2021.

[12] Zhou, Yangming, Yuzhou Yang, Qichao Ying, Zhenxing Qian, and Xinpeng Zhang. "Multimodal fake news detection via clip-guided learning." In *2023 IEEE international conference on multimedia and expo (ICME)*, pp. 2825-2830. IEEE, 2023.

[13] Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794. 2016.

[14] Ali, Armughan, Zeeshan Haider, Hooria Shahbaz, Haris Masood, Bilal Gul, Abdullah Rashid, Zahra Waheed et al. "Towards improved fake news detection using a hybrid RoBERTa and metadata enhanced XGBoost model." *Scientific Reports* (2025).

[15] Deng, Shengli, Yuting Jiang, Hongxiu Li, and Yong Liu. "Who contributes what? Scrutinizing the activity data of 4.2 million Zhihu users via immersion scores." *Information Processing & Management* 57, no. 5 (2020): 102274.

[16] Zhou, Yangming, Yuzhou Yang, Qichao Ying, Zhenxing Qian, and Xinpeng Zhang. "Multi-modal fake news detection on social media via multi-grained information fusion." In *Proceedings of the 2023 ACM international conference on multimedia retrieval*, pp. 343-352. 2023.

[17] Wang, Jingwei, Ziyue Zhu, Chunxiao Liu, Rong Li, and Xin Wu. "LLM-Enhanced multimodal detection of fake news." *PloS one* 19, no. 10 (2024): e0312240.

[18] D'ulizia, Arianna, Maria Chiara Caschera, Fernando Ferri, and Patrizia Grifoni. "Fake news detection: a survey of evaluation datasets." *PeerJ Computer Science* 7 (2021): e518.