International Journal of Current Engineering and Technology
E-ISSN 2277-4106, P-ISSN 2347-5161

Available at https://ijcet.evegenis.org/index.php/ijcet/index

Research Article

Fake News Generation and Detection: Adversarial use of Generative AI in Text Synthesis

Maunik K Shah[1]* and Sarthak Shah[2]

[1,2]Independent researcher, IEEE Senior Member, India

Received 10 Feb 2026, Accepted 12 Mar 2026, Available online 14 Mar 2026, **Vol.16, No.2 (Mar/Apr 2026)**

Abstract

The extensive growth of online media platforms, as well as the extensive use of generative artificial intelligence, has made the production and distribution of fake news particularly more active. Models of languages and vision have now been advanced to create a high level of reality in text, images and videos that become even harder to detect any misinformation by means of the old standard verification channels. The paper is a detailed analysis of fake news and its main premises, the core issues, and the role of generative AI in the development and detection of fake news that keeps changing. The paper examines uni-modal deepfake detection models in text, image and video modalities, and cross-modal consistency analysis, which utilizes semantic consistency between heterogeneous data. In addition, the study contrasts traditional machine learning approaches with more modern ones, detailing the benefits and downsides of each, as well as their suitability for large-scale and multilingual tasks. A number of popular benchmark datasets are presented to give an idea of the experimental evaluation practice. Additionally, stability metrics such as recall (REC), accuracy (ACC), precision (PRE), F1-score (F1), and ROC-AUC are assessed to ascertain the efficacy of detection. Find that deep learning models are more generalizable and perform better overall, but they are still vulnerable to adversarial manipulation and new generative methods.

Keywords: *Fake News Detection, Generative Artificial Intelligence, Deep Learning, Natural Language Processing, Deepfake Detection.*

Introduction

The emergence of digital social media and the accessibility of instant messaging systems have helped modern journalism advance to include sensationalism in news reports. Using literary devices to captivate an audience, even if they distort reality or propagate false information, is never immoral. Some news and social media outlets have a propensity to disregard ethical information gathering and distribution in their pursuit of popularity and likeability in the face of intense competition. The spread of fake news has become one of the burning topics of the modern digital era, where any kind of misinformation can become viral and have a much more significant impact. Studies on fake news became one of the most discussed topics over the last few years, and scholars used different methods to identify any misleading and constructed content [1]. News has a significant part in everyday life and serves as a fortifier of the prejudices and opinions of people.

The field of AI-generated text recognition uses computers with artificial intelligence to find real text and tell it apart from fake or inappropriate text.

*Corresponding author's ORCID ID: 0000-0000-0000-0000
DOI: https://doi.org/10.14741/ijcet/v.16.2.2

The use of deep learning and AI-generated text has grown rapidly in recent years [2]. Nonetheless, the technology has caused several issues, such as incorrect dissemination of information and revealing of privacy. The current emphasis in artificial intelligence is on text detection and identification [3]. The field of NLP emerged concurrently with the start of studies on AI-generated text identification. Models such as RNN, LSTM, and Transformer6 were developed using deep learning technologies, which greatly improved the capabilities of AI-enabled text generation. Any type of high-quality text material can be generated using these models [4]. This includes papers, dialogues, news, and more. They are freely accessible and can be used to transmit dangerous content or mislead consumers through the production of false information.

Artificial intelligence (AI) has the capacity to mimic human procedures, and recent studies have demonstrated that it can identify false news. With the widespread adoption of, the widespread scepticism about AI that was seen in the early years of last decade has been swiftly replaced by excitement. Some of the stuff watch now could be made totally or heavily assisted by AI in the future [5]. There has been an explosion of false information and propaganda on social media in recent times. One may now create fake news with sparse data thanks to recent developments in language models like OpenAI's GPT-2. Generative Adversarial Network (GAN) is one model that can take simple input, like a sentence or a topic, and produce long, legible content; GPT-2 is another model that can do the same for news and fiction articles [6]. Grover, a relatively new model, takes into account domain, date, authors, and headline among other criteria in order to generate fake news using causal language models [7]. Although Grover has proven to be effective, producing relevant news requires a large number of conditional variables. As a means of investigating the issue of MGNs on social media, offers a model for creating synthetic news stories that seem convincing.

Structure of the Paper

The paper structure is as follows: Section I presents the topic of the paper, Section II presents the concept of fake news and challenges, Section III presents the use of generative AI in fake news, Section IV presents the literature review, and Section V wraps up the paper by concluding the research and its future directions.

Fake News: Concepts and Challenges

News stories that are deliberately and demonstrably inaccurate and may cause readers to be misled" is a widely agreed-upon description of fake news. All forms of deceit, including embellishment, partial truth, and concealment, are not captured by this perspective [8]. A relatively new usage of the term "alternative facts" has been to describe manipulations that purposefully include irrelevant material in an attempt to muddy the waters of obvious truth. Misinformation tactics are classified into six categories: (1) out-of-context usage of original sources (2), websites that impersonate someone else; (3), sites that provide explicit false news (4), websites that provide false information (5), websites that modify material and (6), parody. In the context of paltering, which involves deceiving information receivers by making true claims, the misuse of legitimate information is at play.

Uni-modal Deepfake Detection

Figure 1 shows the results of the Uni-modal Deepfake Detection, while Table I provides a summary.

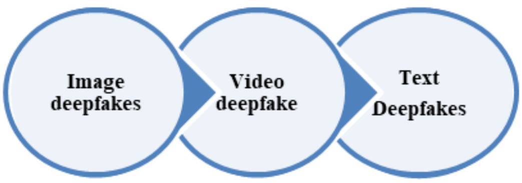

Fig.1 Uni-Modal of Fake News Detection

Image deepfakes: The widespread availability of generative models that generate convincingly faked images—from completely synthetic faces to those with only little adjustments—has shifted the emphasis to image-based deepfake detection [9]. In order to identify discrepancies at the pixel level, earlier approaches depended on spatial pattern learning with CNNs. Their small flexibility to a variety of generative sources, however, led to a transition to frequency-domain analysis with the use of statistical traces to improve generalization.

Video deepfake: Video-based detection is an approach that identifies times when a sequence is being manipulated. Early frame-based techniques were able to identify spatial artifacts, but ignored inter-frame dynamics. 3D CNNs and transformer-based temporal consistency modelling identify the mismatch of lips-synch and motion abnormalities, whereas graph-based modelling detects relational anomalies [10].

Text Deepfakes: Text detection focuses on artificial content such as fake news, which goes beyond a linguistic feature detector to a context-based and domain-sensitive detector. Generalization makes use of multi-task and zero-shot learning, and adversarial fine-tuned models are used to strengthen it [11]. Recent techniques, such as Mc-DNN, process multi-channel text, Obi-LSTM-CNN, and MAGE, address a variety of LLMs [12]. Dealing with aggressive text change and keeping up with evolving LLMs are two of the challenges.

Table 1 Uni-Modal Deepfake Detection Summary

Modality	Key Characteristics & Methods
Image Deep fakes	Detects synthetic or manipulated images using CNN-based spatial artifacts and frequency-domain analysis to capture pixel-level inconsistencies.
Video Deepfakes	Focuses on temporal inconsistencies such as lip-sync errors and motion irregularities using 3D CNNs, transformers, and graph-based models.
Text Deepfakes	Identifies AI-generated fake news and rumors through linguistic, semantic, and contextual modeling with robust and generalizable deep learning methods.

Challenges in Manual and Automated Detection

Some challenges in manual and automated detection. Some of them are:

- Manual methods are very labour-demanding and time-consuming, whereby expert fact-checkers are needed to establish content credibility.
- The sheer amount and speed of proliferation of online news render human-based verification infeasible on scale[13]
- Automated systems of detection cannot effectively deal with linguistic diversity, such as multilingualism, slang and regional terms.
- The changing styles of writing and the pattern of stories make the models that are fixed or rule-based less effective.
- The increased complexity of AI-generated texts makes it very much like that of a human author, and therefore it becomes even harder to spot.

- The fake news is of such a nature because it blends facts and falsities, making it difficult to classify them using binary classifiers.
- The use of adversarial modes like paraphrasing and prompt engineering is in a constant state of development to avoid detection systems.

Generative AI for Fake News Creation

Artificial intelligence systems that are programmed to create fresh and unique media are called "Generative Artificial Intelligence (AI)" systems. Recent developments in AI have greatly reduced the complexity of many common human tasks. One example of these AI technologies is Catgut, which was developed and released by OpenAI in late 2022. Catgut is a conversational agent that functions using text and responds to user queries with text. It has been demonstrated that AI algorithms can effectively identify disinformation and bogus news that could be obstructing optimisation and efficiency [14]. Those who are in favour of utilising AI to identify false news point to a few guidelines that should be followed. These guidelines include software designers creating strategies to counteract fake news, users being able to report false news when they see it, and users being kept informed about how fake news is spreading. DL, ML, and NLP can use text- or image-based information to train models and make better predictions about how true news stories are. In contrast, AI may examine the article's social context by examining the poster's characteristics, like the number of shares or retweets the post received.

Functioning of Generative AI for Fake News Generation

Generative AI algorithms draw on pre-existing datasets to create novel data. Various deep learning architectures, including GANs, VAEs, and Transformers, enable them to operate. Two NNs compete in GANs to produce fresh instances of synthetic data. GANs are used to produce believable images and videos to support fake news that is made up of synthetics [15]. The attention mechanisms used in transformers to produce coherent text sequences. Transformers, similar to GPT models are educated on huge bodies of text. Fake News and other types of text can be created by transformers.

Technical Background

The following section provides a synopsis of the key ideas and technologies that form the basis of the convergence between Generative AI and Fake News.

Generative Artificial Intelligence

A subset of artificial intelligence (AI) known as "generative AI" creates data-like information from scratch. New data samples that cannot be identified as false datasets are generated by training such models.

Generative Adversarial Networks (GANs)

The discriminator and generator, the two parts of a GAN, are both trained to be aggressive [16]. The data samples generated by the generator are intended to be used to deceive the discriminator, whereas the real data is compared to the generated data by the discriminator, and the two models are refined with repeated evaluations.

Variational Autoencoders (VAEs)

Data-generating probabilistic generative networks are VAEs. Their training involves encoding and decoding input data into a latent space, with the goal of producing output samples that are representative of the original sample's distribution.

Transformer Models

Neural network architectures called transformers are built to handle sequential data, especially text. They assess the relevance of various incoming data points using self-attention mechanisms.

BERT: Bidirectional Encoder Representations from Transformers (BERT) is a method that can condition all layers simultaneously on left and right context, allowing it to pre-train deep bidirectional representations from unlabelled text.

GPT: A substantial leap forward in language processing has been made by Generative Pre-trained Transformer (GPT) models, which can now produce coherent and contextually relevant text in response to a prompt [17].

Traditional vs. Deep Machine Learning (to detect fake news) Approaches

Conventional machine learning (ML) methods are very dependent on feature engineering which is mostly manual and statistical learning algorithms. In fake news detection, handcrafted features (e.g. n-grams, TF IDF scores, part of speech tags, sentiment polarity, readability scores and metadata (source credibility, posting time, user behavior) are typically used. Supervised classifiers can be LR, SVM, NB, DTs, or RFs. These are computationally efficient and easier to interpret, and perform reasonably well with small to medium-sized data. Their performance is, however, limited by the quality of features and dependency on domains thus they are not as resistant to changing styles of writing and text written by AI and resembling the pattern of human language.

Deep learning (DL) methods, in contrast, learn hierarchical and contextual representations using raw text and do not require manual feature extraction. Semantic meaning, long-range dependencies, and contextual nuances that help identify advanced fake news are modelled by CNNs, RNNs, LSTMs, Bi-LSTMs, transformer-based models (e.g., BERT and GPT-style encoders)[18]. Deep learning models typically have higher ACC and generalization performance, especially when applied to large-scale and multilingual data. They are however highly computationally intensive and require large labeled data, which makes them difficult to interpret, which makes them difficult to explain and deploy in the real world.

Datasets

Run tests on two news datasets with publisher metadata:

- In the CNN/Daily Mail archive, find more than 300,000 original news articles and highlighted passages [19]. Staff writers from CNN and the UK's Daily Mail penned it. The 3 sets—training, validation, and testing—form the official split.
- The All the News dataset contains 143,000 articles and commentaries published by 15 different American publishers. To make sure that each publisher's news has enough news to display the matching template patterns, selected data from the top five common publishers (NPR, NY Post, Reuters, Washington Post, and Breitbart).
- Fake Newsnet dataset to teach model and make GPT-2 work better. Here is news data X compiled from two sources: PolitiFact and Gossip Cop [20]. With an emphasis on fact-checking, Gossip Cop covers celebrity news. Similar to this, PolitiFact verifies the accuracy of political news and stories. False or true news stories are categorised in this dataset. While Gossip Cop has 3,586 genuine stories and 2,230 false stories, PolitiFact has 2,645 true stories and 2,770 false stories.
- LIAR the information utilized for this research came from over 12,800 statements that were manually annotated and found on the PolitiFact website. From completely fake to completely true, these claims have been sorted into six groups: "pants on fire," "false," "barely true," "half true," "mostly true," and "true." Textual assertions and information, including speaker, subject, and context, are both included in the collection [21].

Feature Extraction

Table II shows that fake news can be broken down into four types of features that are meant to do harm: confuse people, change their minds, get people's attention, and other general features [22]. Each of the other three kinds displays unique characteristics that mirror the unique motivations behind their development, with the exception of the overarching characteristics. Give some examples of each in Figure 2, with the red-coloured feature words or symbols matching to them; its talk about the distinctions then:

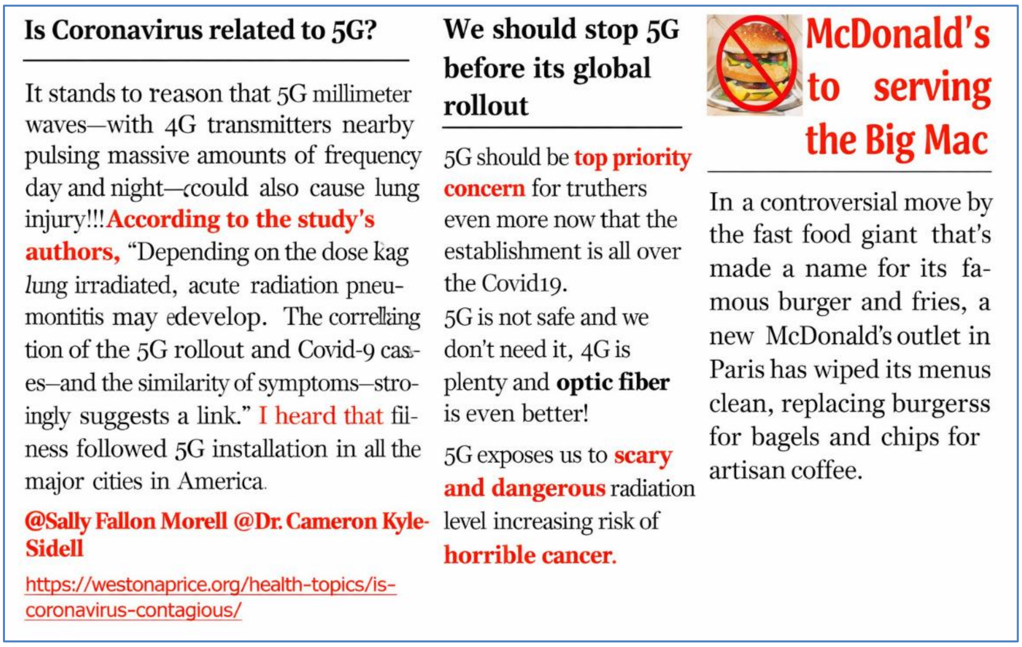

Fig.2 Examples of fake news that aim to mislead in three distinct ways: the (a) is deceiving the general population, the (b) is swaying public opinion, and the (c) is drawing in users

Mislead the Public: The goal of spreading false information through social media is to deceive others into thinking the news stories are factual. Therefore, the content of the text is quite similar to actual news, with the exception of a few minor modifications in the use of special symbols to bolster the news's credibility, as illustrated in Figure 2a. These symbols include personal pronouns, URLs, and "@" tags, among others.

Manipulate Opinions: This type of false news primarily aims to sway people's opinions in favour of its own viewpoints by manipulating their opinions. This sort of fake news is distinct from the "Mislead the

Public" category since it does not aim to make the news seem more credible.
Attract User Attention: The goal of this sort of disinformation is to generate more interest, traffic, or clicks. Figure. 2c shows that this sort of fake news differs from the previous two in that it depends mostly on headlines and cover page pictures; as a result, it frequently uses trendy subjects, appealing images, or clickbait.

Table 2 Feature Extraction Based on Creation Intentions of Fake News

Intention Type	Purpose	Representative Features
Mislead the Public	Make fake content appear truthful and credible	Real-news writing style, personal pronouns, URLs, mentions (@), formal tone
Manipulate Opinions	Influence readers' beliefs or attitudes	Emotionally charged words, biased language, subjective expressions, persuasive cues
Attract User Attention	Increase clicks, traffic, or virality	Clickbait headlines, sensational phrases, hot topics, exaggerated claims, visuals
General Features	Common across all fake news types	Text length, word frequency, punctuation patterns, syntactic structure

Cross-Modal Consistency Analysis

Cross-modal consistency analysis verifies whether different content modalities (text, images, and videos) align. Many fake news pieces use authentic images with misleading text, creating a false narrative without modifying the image.
Text-image mismatch Detection: AI models compare headlines, captions, and textual content with accompanying images to detect inconsistencies. For example, a misleading headline about a protest may use an unconnected image from a different event.
Visual-Text Semantic Verification: NLP models analyze image descriptions and compare them with textual claims, checking whether the semantic meaning aligns.
Video-Text Alignment Analysis: AI models analyze video transcripts and verify whether speech content aligns with visual evidence, detecting fabricated or edited content.
The cross-modal analysis is adequate but limited when attackers subtly modify narratives without creating apparent contradictions.

Evaluation Metrics

The efficacy of models for detecting false news is evaluated using a handful of commonly used performance metrics:

Accuracy

The percentage of examples that have been correctly classified out of all the predictions is measured by ACC, which is a measure of the general validity of the model [23]. Although it gives a general overview of the performance, ACC may be false in unbalanced datasets where fake news samples are much less or much more real news.

Precision

PRE denotes reliability of positive predictions and considers the ratio of the number of correctly identified fake news to all the instances called as fake. The low level of PRE means that the rate of false positives is very low, and this is essential in order to prevent the wrong definition of real news as fake.

Recall

REC assesses the REC capacity of the model with the aim of measuring the occurrence of correctly predicted fake news of all the true fake news samples. A high level of REC is an indicator of successful detection with a low false negative rate.

F1-score

A balanced evaluation metric, the F1 takes a harmonic average of REC and PRE and is useful in cases of class imbalance. ACC and REC are both compromised in this case.

ROC-AUC

ROC-AUC is used to estimate the capacity of the model to distinguish fake news and genuine news at various classification levels. A greater ROC-AUC value suggests greater capability and strength of discrimination of the detection model

Literature Review

The following section identifies the related works on the Fake News Generation and Detection Use of Generative AI in Text.

Nandan et al. (2025) developed this challenge, using various models of deep learning- LSTM, ALBERT, FNNet, and CNN+RNN model. These models are then compared on their ACC to determine the most appropriate architecture for detecting fake news. The LSTM model has recorded an ACC of 96.25 in textual data where the temporal dependencies were also modelled. Using ALBERT, best suited to NLP tasks scored 100% indicating that it is capable of identifying intricate patterns of a language. Since the architecture of FNNet was developed to identify fake news, it reached almost 93% of ACC in its work. Finally, hybrid CNN + RNN, a combination of spatial and temporal feature extraction, which had 99% ACC [24].

Wang and Long (2025) propose a new Global-Local News Detection Model that uses a combination of

BERT, BiLSTM networks, Text CNN and attention networks to improve the detection of fake news that is created by AI. New data, which was created with the help of GPT-4 and included 42 news categories, has been created to act as an exhaustive and diverse base to train and test the model. The experimental findings suggest that the proposed model has an ACC and F1 of 0.82 which is higher than conventional methods [25].

Abburi et al. (2024) presented a neural model that is fine-tuning-based and a combination of transformer models, linguistic features, and state-of-the-art embedding models. additionally maintained a training dataset of various samples across domains and models of language to refine pretrained language models. Compared the work of model with the state-of-the-art methods, and the comparative analysis illustrates that model is always the best in the choice of different data sets with the help of the accepted evaluation criteria [26].

Grecu and Breaban (2024) introduce a two-sided AI-generated text detector, which uses both conventional machine learning (ML) methods and novel fine-tuned large language models (LLMs). On a detailed dataset that includes more than 350,000 samples provided by five benchmark sources, the model has shown good performance, and traditional ML algorithms attain ACC of 91-92% (0.97 ROC-AUC) and fine-tuned LLMs like BERT and Roberta with 97-98% ACC (0.99 ROC-AUC). Tritest, a working tool, is an application to bring together these models to give a more detailed analysis of AI-generated content, including the probability at the paragraph level [27].

Singh and Namin (2024) use a dual architecture of large language models and a task-specific fake news detector, that is, embedding-based retrieval with sentence transformers and Facebook AI Similarity Search (FAISS), as well as support generation with GPT-Neo, Roberta, and personalized RAG architectures. Additionally, employ a passive-aggressive classifier trained on "Fake" and "Real" datasets from a public GitHub repository to assess the likelihood of generated responses being classified as "Fake" or "Real." This pipeline evaluates the authenticity of news articles and incorporates believability scores to enhance interpretability. Results indicate that while all models perform comparably, the custom RAG model consistently excels in providing contextually grounded and highly relevant fake information, in these fake news scenarios[28].

Walambe et al. (2023) These types of automatically created content can be identified with the use of ML and DL. Gathering and labelling massive amounts of text data, including both real and fraudulent news pieces, headlines, and data from any other source, is a common practice. Thus, issues with human bias, label ACC, and data quality from aggregated sources persist in existing systems. Here introduce and showcase a GAN-based textual deep fake detection and generating system [29].

Table III provides a summary of Fake News Generation and Detection Generative AI in Text, including study, Approach, Key Findings and challenges.

Table 3 Comparative Analysis of Fake News and AI-Generated Text Detection Approaches

Reference	Study On	Approach	Key Findings	Challenges / Limitations	Future Directions
Nandan et al. (2025)	Fake news detection using deep learning architectures	LSTM, ALBERT, FNNet, CNN+RNN hybrid	ALBERT achieved 100% accuracy, CNN+RNN reached 99%, LSTM captured temporal dependencies with 96.25%, FNNet achieved ~93%	Risk of overfitting with extremely high accuracy; limited discussion on dataset diversity and real-world generalization	Validation on larger, cross-domain datasets; robustness testing against adversarial and unseen AI-generated content
Wang and Long (2025)	AI-generated fake news detection	Global–Local News Detection Model using BERT, BiLSTM, TextCNN, attention	Achieved 0.82 accuracy and F1-score, outperforming traditional models; effective global and local feature fusion	Performance lower than pure transformer-based models; computational complexity	Scaling to multilingual datasets; improving efficiency and accuracy with lightweight transformers
Abburi et al. (2024)	Detection of LLM-generated fake news	Fine-tuned transformer models combined with linguistic features and embeddings	Consistently outperformed state-of-the-art methods across multiple datasets	High dependency on curated datasets; fine-tuning cost and scalability issues	Automated dataset expansion; continual learning for evolving LLM-generated text
Grecu and Breaban (2024)	AI-generated text detection	Dual approach: traditional ML + fine-tuned LLMs (BERT, RoBERTa); TruAIText tool	ML models achieved 91–92% accuracy, LLMs reached 97–98% accuracy with 0.99 ROC-AUC; paragraph-level explainability	High computational requirements; potential bias toward benchmark datasets	Explainable AI enhancements; deployment-friendly lightweight detection systems
Singh and Namin (2024)	Fake news detection and generation with retrieval pipelines	LLMs, sentence transformers, FAISS, GPT-Neo, RoBERTa, custom RAG, Passive Aggressive Classifier	Custom RAG model excelled in contextual relevance; believability scores improved interpretability	Comparable classification performance across models; generation-focused bias	Stronger factual verification modules; tighter integration of detection and fact-checking
Walambe et al. (2023)	Textual deepfake detection and generation	GAN-based framework for text generation and detection	Demonstrated feasibility of GANs for detecting auto-generated fake text	Data labeling bias; instability in GAN training; limited scalability	Hybrid GAN–transformer models; improved data curation and adversarial robustness

Conclusion and Future Work

The security of online data ecosystems has emerged as a pressing societal concern in the era of pervasive social media and generative AI. Fake news was covered extensively in the study, from its conceptual basis to the issue of detecting it and the growing importance of generative AI in both processes. Examining text, image, and video models' uni- and cross-modal detection strategies, the study found that deep learning has advanced significantly, particularly in transformer-based models, which have improved the ability to detect subtle linguistic, semantic, and contextual features associated with false information.

Simultaneously, the survey reiterated that the growing realism of AI-generated content, adversarial manipulation approaches, and the integration of factual and fake information still persist in the reduction of the strength of the current detection systems. The next generation of fake news detection models that are more resilient and explainable and can adapt to the fast-changing generative models should be a major focus of future research. This involves investigation of self-supervised and lifelong learning, self-supervised multimodal fusion with more powerful cross-modal reasoning and low-resource multilingual detection to counteract global misinformation.

Furthermore, the verification of human-in-the-loop, the ontology of social context, and propagation patterns can contribute to the rise in the reliability. Moral principles, disclosure, and legal compliance will also be very important in the responsible implementation of such systems. In general, long-term interdisciplinary cooperation is necessary to curb the effects of AI-based fake news, as well as to protect the population's trust in online media.

References

[1] D. Trandabăț and D. Gifu, “Discriminating AI-generated Fake News,” Procedia Comput. Sci., vol. 225, pp. 3822–3831, 2023, doi: 10.1016/j.procs.2023.10.378.

[2] R. Palwe and A. Kumar, “Redefining usability in the age of generative AI: Towards a new evaluation paradigm,” Int. J. Comput. Artif. Intell., vol. 6, no. 2, pp. 155–163, Jul. 2025, doi: 10.33545/27076571.2025.v6.i2b.193.

[3] Y. Hui, “Using generative adversarial network to improve the accuracy of detecting AI-generated tweets,” Sci. Rep., vol. 14, no. 1, p. 29322, Nov. 2024, doi: 10.1038/s41598-024-78601-1.

[4] V. Shah, “Managing Security and Privacy in Cloud Frameworks : A Risk with Compliance Perspective for Enterprises,” Int. J. Curr. Eng. Technol., vol. 12, no. 6, pp. 606–618, 2022, doi: 10.14741/ijcet/v.12.6.16.

[5] H. P. Kapadia and K. B. Thakkar, “Generative AI for Real-Time Customer Support Content Creation,” J. Emerg. Technol. Innov. Res., vol. 10, no. 12, pp. 36–43, 2023.

[6] V. Verma, “Deep Learning-Based Fraud Detection in Financial Transactions : A Case Study Using Real-Time Data Streams,” vol. 3, no. 4, pp. 149–157, 2023, doi: 10.56472/25832646/JETA-V3I8P117.

[7] S. Thangavel, “AI Enhanced Image Processing System For Cyber Security Threat Analysis,” 2024.

[8] K. M. DSouza and A. M. French, “Fake news detection using machine learning: an adversarial collaboration approach,” Internet Res., vol. 34, no. 5, pp. 1664–1678, 2024, doi: 10.1108/INTR-03-2022-0176.

[9] N. Khan, T. Nguyen, A. Bermak, and I. Khalil, “Unmasking Synthetic Realities in Generative AI: A Comprehensive Review of Adversarially Robust Deepfake Detection Systems,” arXiv, 2025.

[10] X. Yu et al., “Fake Artificial Intelligence Generated Contents (FAIGC): A Survey of Theories, Detection Methods, and Opportunities,” arXiv, vol. 3162, pp. 0–3, 2024.

[11] A. Dudhipala, R. Karne, and P. K. Pativada, “Prompt2Graph: Leveraging LLMs to Construct Knowledge Graphs from Technical Manuals,” in 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), IEEE, Sep. 2025, pp. 912–919. doi: 10.1109/ICIMIA67127.2025.11200177.

[12] V. Rajendran, D. Besiahgari, S. C. Patil, M. Chandrashekaraiah, and V. Challagulla, “A Multi-Agent LLM Environment for Software Design and Refactoring: A Conceptual Framework,” in SoutheastCon 2025, IEEE, Mar. 2025, pp. 488–493. doi: 10.1109/SoutheastCon56624.2025.10971563.

[13] E. Lamprou, N. Antonopoulos, I. Anomeritou, and C. Apostolou, “Characteristics of Fake News and Misinformation in Greece: The Rise of New Crowdsourcing-Based Journalistic Fact-Checking Models,” Journal. Media, vol. 2, no. 3, pp. 417–439, Jul. 2021, doi: 10.3390/journalmedia2030025.

[14] R. Raman et al., “Fake news research trends, linkages to generative artificial intelligence and sustainable development goals,” Heliyon, vol. 10, no. 3, p. e24727, Feb. 2024, doi: 10.1016/j.heliyon.2024.e24727.

[15] A. Loth, M. Kappes, and M.-O. Pahl, “Blessing or curse? A survey on the Impact of Generative AI on Fake News,” arXiv, 2024.

[16] S. K. Tiwari, “Automating Behavior-Driven Development with Generative AI: Enhancing Efficiency in Test Automation,” Front. Emerg. Comput. Sci. Inf. Technol., vol. 02, no. 12, pp. 01–14, Dec. 2025, doi: 10.64917/fecsit/Volume02Issue12-01.

[17] S. B. Karri, S. Gawali, S. Rayankula, and P. Vankadara, “AI Chatbots in Banking: Transforming Customer Service and Operational Efficiency,” in Advancements in Smart Innovations, Intelligent Systems, and Technologies, 2025, pp. 61–81. doi: 10.3233/FAIA251498.

[18] S. Amrale, “Anomaly Identification in Real-Time for Predictive Analytics in IoT Sensor Networks using Deep,” Int. J. Curr. Eng. Technol., vol. 14, no. 6, pp. 526–532, 2024, doi: 10.14741/ijcet/v.14.6.15.

[19] W.-Y. Wang, Y.-C. Chang, and W.-C. Peng, “Style-News: Incorporating Stylized News Generation and Adversarial Verification for Neural Fake News Detection,” in Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers), 2024, pp. 1531–1541. doi: 10.18653/v1/2024.eacl-long.92.

[20] A. Mosallanezhad, K. Shu, and H. Liu, “Topic-Preserving Synthetic News Generation: An Adversarial Deep Reinforcement Learning Approach,” arXiv, no. 2, 2020.

[21] D. Patel, “AI-Enhanced Natural Language Processing for Improving Web Page Classification Accuracy,” vol. 4, no. 1, pp. 133–140, 2024, doi: 10.56472/25832646/JETA-V4I1P119.

[22] B. Hu, Z. Mao, and Y. Zhang, “An overview of fake news detection: From a new perspective,” Fundam. Res., vol. 5, no. 1, pp. 332–346, Jan. 2025, doi: 10.1016/j.fmre.2024.01.017.

[23] N. K. Prajapati, “Federated Learning for Privacy-Preserving Cybersecurity: A Review on Secure Threat Detection,” Int. J. Adv. Res. Sci. Commun. Technol., vol. 5, no. 4, pp. 520–528, Apr. 2025, doi: 10.48175/IJARSCT-25168.

[24] K. P. Nandan, B. Pakruddin, S. Afridi, S. K. R., and S. V. Pati, "Real-Time Detection of Fake News Articles Using Deep Learning Techniques," in 2025 International Conference on Next Generation Communication & Information Processing (INCIP), IEEE, Jan. 2025, pp. 687–691. doi: 10.1109/INCIP64058.2025.11019208.
[25] Y. Wang and W. Long, "Global-Local Ensemble Detector for AI-Generated Fake News," IEEE Access, vol. 13, pp. 69779–69789, 2025, doi: 10.1109/ACCESS.2025.3562154.
[26] H. Abburi, N. Pudota, B. Veeramani, E. Bowen, and S. Bhattacharya, "Toward Robust Generative AI Text Detection: Generalizable Neural Model," in 2024 International Conference on Machine Learning and Applications (ICMLA), IEEE, Dec. 2024, pp. 1651–1656. doi: 10.1109/ICMLA61862.2024.00255.
[27] C.-S. Grecu and M.-E. Breaban, "A Dual-Approach for AI-Generated Text Detection," in 2024 26th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), IEEE, Sep. 2024, pp. 206–214. doi: 10.1109/SYNASC65383.2024.00043.
[28] S. Singh and A. S. Namin, "Adversarial Training of Retrieval Augmented Generation to Generate Believable Fake News," in 2024 IEEE International Conference on Big Data (BigData), IEEE, Dec. 2024, pp. 3589–3598. doi: 10.1109/BigData62323.2024.10825933.
[29] R. Walambe, P. Chaudhary, A. Bajaj, A. S. Rathore, V. Jain, and K. Kotecha, "Generative Adversarial Networks for Mitigating Bias in Disinformation," in 2023 IEEE International Conference on Contemporary Computing and Communications (InC4), IEEE, Apr. 2023, pp. 1–6. doi: 10.1109/InC457730.2023.10262880.