

A Systematic Review of Deep Learning Approaches for Fake News Detection

Mosimabale Agbabiaka ¹; Emeka Ogbuju ²; and Francisca Oladipo ³

¹Department of Computer Science, Federal University Lokoja, Abuja, Nigeria. ²Department of Computer Science, Miva Open University, Abuja, Nigeria. ³Department of Computer Science, Thomas Adewunmi University, Oko, Nigeria.

Corresponding author: mosimabale.agbabiaka@fulokoja.edu.ng

DOI: <https://doi.org/10.62154/ajastr.2025.021.01011>

Abstract

Distinguishing real news from fake news on online platforms remains a critical research challenge. This systematic literature review (SLR) evaluates hybrid and non-hybrid deep learning (DL) approaches for fake news detection across ISOT, WELFake, and FA-KES datasets, comparing their mean accuracy and model performance. PRISMA guidelines were followed to search databases such as Google Scholar, Scopus, IEEE Xplore, and MDPI using search terms such as "fake news detection", "deep learning", and "hybrid models" for articles published between 2018 and 2025. Out of 154 articles surveyed, 34 duplicates were removed, 70 were excluded for irrelevance, and 50 were included. Addressing three research questions shows that hybrid models like Convolutional Neural Networks and Long Short-Term Memory (CNN-LSTM) consistently outperform non-hybrid models (e.g., CNN) by 0.16% – 8.6% accuracy across the three datasets (mean accuracy: hybrid 85.46%; non-hybrid 82.46%). Combining the right models in a hybrid approach improves performance, yielding more reliable results. Limitations include dataset biases. Future research could explore multimodal datasets and BERT base models to improve robustness.

Keywords: Fake News, Deep Learning, Hybrid Models, Detection Accuracy, Text Classification.

Introduction

Fake news refers to intentionally fabricated content designed to deceive readers while mimicking legitimate reporting (Allcott & Gentzkow, 2017). Social media accelerates the dissemination of fake news compared to true news, enabling quicker reach than trustworthy news (Vosoughi et al., 2018). This has real-world impacts, posing risks to societal stability, such as swaying elections (Allcott & Gentzkow, 2017), exacerbating conflicts, as seen in Nigeria's farmer-herder clashes, where misrepresented images fueled violence that resulted in loss of lives (Adegoke et al., 2018). Moreover, the proliferation of fake news poses significant threats that extend beyond conflict scenarios, impacting critical domains such as healthcare. A notable example is a case in China where a false advertisement for cancer treatment resulted in a patient's death (Alghamdi et al., 2023). During the COVID-19 pandemic, misinformation on preventing and curing the pandemic

led to hundreds of deaths and thousands of hospitalizations (Coleman, 2020). Ogbuju et al. (2023) highlight the threat of fake news to a society and individuals, underscoring the need for effective detection mechanisms.

These impacts underscore the need for ongoing research in fake news detection, including the development of sophisticated, automated detection systems that can efficiently identify fake news or misinformation, thereby advancing the field of fake news detection. DL models excel in processing textual data and extracting relevant features for fake news identification (Kaliyar et al., 2020). However, hybrid approaches combining multiple architectures remain underexplored. This research systematically reviews DL approaches for fake news detection, focusing on hybrid and non-hybrid approaches, comparing their performance on standard datasets such as ISOT, WELFake, and FA-KES, helping identify their suitability for real-world deployment.

Problem Statement

Despite advances in DL for text classification, existing reviews lack in-depth analysis or systematic comparisons of hybrid and non-hybrid DL approaches across standard datasets such as ISOT, WELFake, and FA-KES, where data size and context pose challenges (Mridha et al., 2021). While prior surveys (Hu et al., 2022; Mridha et al., 2021) explored DL for fake news, none systematically compared hybrid and non-hybrid models across standard datasets using PRISMA methodology. This gap hinders the understanding of how model combinations enhance detection accuracy, particularly in diverse data contexts, motivating this SLR to address these deficiencies using the PRISMA methodology.

Research Methodology

This SLR follows PRISMA guidelines for transparency in article selection for this review (Page et al., 2021). Peer-reviewed articles on fake news detection using a DL approach (2018-2025) were targeted to capture recent advancements. Articles were selected based on their relevance to fake news detection, DL, and hybrid models and datasets.

Database Source and Search Procedure

A total of 154 articles were collected from notable sources, including Google Scholar, Scopus, MDPI, and IEEE Xplore, using the search query: ("fake news" AND ("detection" OR "classification")) AND ("deep learning" OR "hybrid models"). After removing 34 duplicates, 70 articles were excluded for irrelevance. Leaving 50 articles for analysis (see Fig. 1: PRISMA Chart).

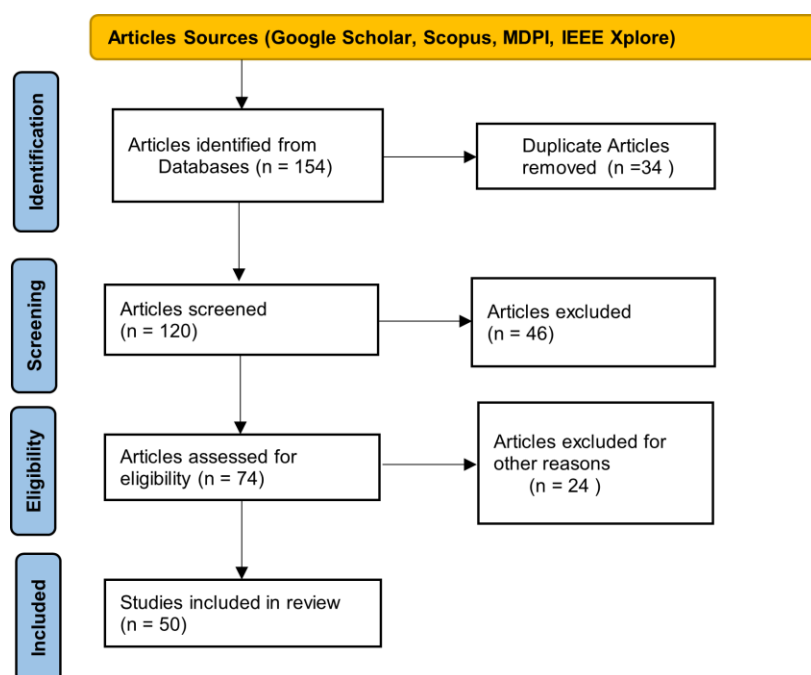


Fig. 1: PRISMA Chart for Article Selection

There may also be bias toward studies reporting high accuracy, as journals tend to favour significant results, potentially skewing performance and overestimating their effectiveness in real-world scenarios.

Inclusion and Exclusion Criteria

The research articles included in this review meet the following criteria:

- Written in English
- Relevant to fake news detection
- Relevant to deep learning
- Relevant to hybrid deep learning
- Published between 2018 and 2025
- Peer-reviewed articles

The articles that do not meet the above criteria were excluded.

Data Extraction and Synthesis

This study adhered to the PRISMA guidelines (Page et al., 2021) in selecting articles for this research, which involved defining research questions, developing a search query, and applying inclusion and exclusion criteria. From 154 articles collected across Google Scholar, Scopus, IEEE Xplore, and MDPI, 34 duplicates were removed. Fifty (50) articles were included in the research after excluding 70 articles that were not relevant or fell short of the inclusion criteria. Data were extracted narratively, focusing on model types, datasets (ISOT, WELFake, FA-KES), accuracy, and key findings, followed by quantitative aggregation of mean accuracy. Table 1 gives the details of the included studies.

Author & Year	Model	Dataset	Accuracy (%)	Key Findings
(Verma et al., 2021)	CNN	WELFake	93.79	Introduced the WELFake benchmark dataset; CNN achieved baseline performance.
(Kaliyar et al., 2020)	FNDNet (CNN)	Kaggle Fake News	98.36	FNDNet outperformed baselines by 5–10% using GloVe embeddings and deep CNN layers.
(Goldani et al., 2021)	CNN	ISOT & LIAR	7.9 improvement on ISOT	Non-static embeddings improved accuracy by 7.9% on ISOT and 2.1% on LIAR vs. static embeddings.
(Deepak & Chitturi, 2020)	LSTM	George McIntire	91.3	Live mining of auxiliary features (author, source) improved performance by 7.6%.
(Sastrawan et al., 2022)	CNN, BiLSTM, ResNet	ISOT	99.95	Data augmentation mitigated class imbalance, boosting BiLSTM performance.
(Padalko et al., 2024)	BiLSTM, BiLSTM+Attn	WELFake	97.49, 97.66	The attention-based BiLSTM outperformed the plain BiLSTM by 0.17%.
(Nasir et al., 2021)	CNN-RNN	ISOT, FA-KES	99, 60	The hybrid CNN-RNN outperformed the standalone CNN and RNN on both datasets.
(Ouassil et al., 2022)	CNN-BiLSTM	WELFake	97.74	Hybrid embeddings (Word2Vec CBOW + Skip-Gram) improved accuracy over single embeddings.
(Tokpa et al., 2023)	CNN-BiLSTM CNN-LSTM	FA-KES, ISOT	68, 99 63, 99	Hybrid models outperformed non-hybrid and ML baselines, especially on FA-KES.
(Mahara et al., 2023)	CNN-BiLSTM, CNN-LSTM	Fake News Healthcare (FNH)	96.94, 97.09 (F1-score)	Feature-based models outperformed content-only models, and augmentation improved the balance.
(Jian et al., 2024)	BiLSTM+Attn	ISOT	99.98	An attention mechanism with the ADAM optimizer achieved near-perfect classification accuracy.

(Umar et al., 2023)	CNN-GRU	ISOT FA-KES WELFake	99 59 91	CNN-GRU achieved high accuracy with reduced computational complexity vs. LSTM.
(Saleh et al., 2021)	CNN	ISOT FA-KES	99.99 53.99	Hyperparameter optimization maximized CNN performance; FA-KES posed generalization challenge.
(Sarin & Kumar, 2025)	CNN+BiRNN CNN+BiLSTM CNN+BiGRU	ISOT, WELFake	99.66, 95.53, 95.86	Context-aware hybrids used fewer parameters; ISOT showed superior robustness.

Table 1. Summary of Included Studies

This review aims to:

- 1) identify commonly used DL models for fake news detection,
- 2) evaluate hybrid model architectures, and
- 3) compare their performance across datasets.

This research aims to answer the following research questions:

- RQ1: What DL models are commonly used for fake news detection, and what are their performances?
- RQ2: What hybrid DL models are used for fake news detection, and what are their performances?
- RQ3: How do non-hybrid and hybrid models compare on benchmarked datasets?

Results

This section analyzes the results of relevant studies, addressing the research question, by presenting and examining performance across ISOT, FA-KES, and WELFake datasets for hybrid and non-hybrid models, comparing their performance.

Dataset Descriptions

- 1 **ISOT Dataset:** Comprises 23,481 fake news and 21,417 real news articles collected from real-world sources on world news, politics and government news. The real news was gathered from reuters.com, while the fake news was sourced from unreliable websites (Ahmed & Traore, 2017).
- 2 **FA-KES Dataset:** A smaller and domain-specific dataset focusing on news from the Syrian war. It contains 804 news articles (426 true, 378 false) (Abu Salem et al., 2019). Its limited size and contextual complexity pose challenges, resulting in lower accuracy (Nasir et al., 2021; Saleh et al., 2021; Umar et al., 2023).
- 3 **WELFake Dataset:** Contains over 72000 news articles, with 35,028 real news and 37,206 fake news. It is a combination of four datasets (Kaggle, McTntire, Reuters,

and Buzzfeed), designed to reduce overfitting. It yields high accuracy owing to its large size (Hashmi et al., 2024).

Rq1. Common dl models used for fake news detection and their performance?

Deep learning has continued to advance the field of Natural Language Processing (NLP), including tasks such as fake news detection (Nasir et al., 2021; Ouassil et al., 2022). It can learn complex patterns in data, enabling it to identify subtle characteristics and nuances in news content that may indicate the presence of fake news. Studies suggest that various deep learning architectures, such as Recurrent Neural Networks (RNNs) and their variants like Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Bidirectional LSTM (BiLSTM), mitigate the vanishing gradients limitation in RNNs (Noh, 2021) and effectively extract relevant sequential features (Nasir et al., 2021). This makes them efficient for fake news detection (Li & Lei, 2022). Convolutional Neural Networks (CNNs) have also been utilized for text classification, leveraging their ability to identify local patterns and relationships within a text (Kaliyar et al., 2020), focusing on individual words and their context. CNNs are highly adaptable and excel in extracting local patterns from data (Liang et al., 2017). A range of approaches and architectures has been investigated, with a focus on deep learning.

CNN-based models excel at capturing n-gram-like features and syntactic patterns, enabling them to achieve high accuracy in prediction. Kaliyar et al. (2020) proposed a novel deep CNN model for detecting fake news using GloVe embedding. The model (FNDNet) achieved 98.36% accuracy on the test data, improving upon baseline models by 5-10%. Goldani et al. (2021) proposed a CNN model comparing static and non-static embeddings. In non-static embeddings, updates occur during training, which improves the model's accuracy by 7.9% on ISOT and 2.1% on LIAR. Otor et al. (2023) also presented RNN and CNN models on a novel dataset extracted from Nigerian social media, with RNN achieving 82.34% and CNN achieving 81.96%. CNN had a better recall of 98.03% compared to RNN's 50.61%, but the model shows signs of overfitting. Sastrawan et al. (2022) proposed models (CNN, BiLSTM, and ResNet) were trained on four datasets: ISOT, Fake News, Fake or Real, and the Fake News detection dataset. Class imbalance was addressed through data augmentation, utilizing back-translation. BiLSTM had the best performance across the four datasets in all metrics. Particularly scoring 99.95% across metrics on the ISOT dataset, highlighting the importance of augmentation in the model's performance.

Data preprocessing has a significant impact on the performance of machine learning models, as it converts raw or complex data into a simplified, informative format, thereby reducing dimensionality, noise, and redundancy, which can lead to better accuracy and increased computational efficiency (Uniyal, 2024). Verma et al. (2021) created a novel benchmark dataset (WELFake dataset), designed to overcome overfitting. The model combined linguistic features with word embeddings to improve the model's performance. Kausar et al. (2022) employed TD-IDF and N-gram for better feature extraction and representation, using the WELFake and Kaggle FakeNews datasets for the LSTM and BERT

models. LSTM achieved 96% accuracy and 95.5% precision for the WELFake dataset. Performance on the Kaggle Fake News dataset was slightly lower, with 94.6% accuracy and 94.6% precision. BERT slightly outperformed the LSTM model with 0.8% accuracy. Ali et al. (2023) utilized statistical techniques, principal component analysis (PCA) for dimensionality reduction with the BiLSTM model, achieving 98.52% accuracy when trained on the WELFake dataset. Augmentation via back-translation and class balancing with SMOTE (Esan et al., 2023) mitigates imbalance, boosting LSTM performance by up to +5.36% on skewed distributions (Sastrawan et al., 2022). Auxiliary feature integration (e.g., author metadata, source domain) further enhances prediction accuracy observed in content-augmented models (Deepak & Chitturi, 2020).

RNN variants, like LSTM and BiLSTM, excel in modelling long-range dependencies and benefit from attention mechanisms, which improve precision by dynamically weighting salient tokens to improve models' efficiency (Jiang et al., 2025). Ramya & Eswari (2022) study demonstrates that the attention mechanism is key to enhancing the performance of models. Jian et al. (2024) introduced a novel model using self-attention with BiLSTM with ADAM optimizers, achieving 99.98% accuracy on the ISOT dataset. Padalko et al. (2024) also implemented an attention-augmented model showing marginal gains (+0.17%) over BiLSTM with no attention, validating their role in enhancing contextual focus.

Deep learning models have also demonstrated superior generalization and scalability (Jaybhaye et al., 2023; Chang, 2024), while ML classifiers (e.g., Logistic Regression, Passive Aggressive) achieve competitive results on structured features (Alarfaj & Khan, 2023). They overfit on imbalanced or high-dimensional text. DL maintains robustness across dataset complexity. Hashmi et al. (2024) presented a comprehensive framework for fake news detection by experimenting with ML, DL, and transformer models across three distinct datasets (WELFake, FakeNewsNet, and FakeNewsPrediction). Regularization and hyperparameter techniques were employed. DL models generally maintained consistency and generalized well across the three datasets, indicating their ability to classify varying levels of complexity. The transformer models showed varied performance and required more computational time than ML and DL models. Mahara et al. (2023) addresses health-related misinformation using ML and hybrid DL models. Sentiment scores and source metadata were extracted from the dataset to enhance classification. CNN-LSTM and CNN-BiLSTM achieved 97.09% and 96.94% F1-score, respectively, demonstrating DL's effectiveness in classifying health news.

Non-hybrid DL models demonstrate good performance in fake news prediction, achieving mean accuracies of 99.11% (ISOT), 52.11% (FA-KES), and 96.17% (WELFake) (Table 8). This highlights the strong efficacy on large-scale, general-domain data but reveals vulnerability in low-resource, domain-specific contexts, a gap addressed in RQ2 and RQ3 through hybrid architectures.

RQ2. Hybrid DL models used for fake news detection, and their performance?

Researchers have explored hybrid DL models that combine different models, such as CNN and RNN or their variants, to integrate architectures for better feature capture (Mridha et al., 2021), to learn more comprehensive and unified features from input texts (Nasir et al., 2021). A hybrid DL model is a neural network that combines components, leveraging their strengths to achieve improved performance. According to Gomedé, (2024) and Tokpa et al. (2023) A hybrid DL architecture can address the limitations of non-hybrid models in fake news prediction, as demonstrated by various studies, especially on benchmark datasets such as ISOT, FA-KES, and WELFake.

With appropriate feature integration and extraction, CNNs combined with sequential models, such as RNN and its variants, can leverage local pattern extraction and sequential context modelling to address the limitations of non-hybrid models. Nasir et al. (2021) proposed a hybrid of CNN and RNN on the ISOT and FA-KES datasets and achieved a better result on both datasets, scoring 99% and 60% accuracy compared to CNN's 99% and 50% and RNN's 98% and 50%. Tokpa et al. (2023) employed a hybrid model combining CNN and BiLSTM on the ISOT and FA-KES datasets, outperforming non-hybrid models such as CNN, LSTM, and BiLSTM. Kumar et al. (2020) implemented an attention-augmented CNN+BiLSTM with a 200-dimensional GloVe embedding to capture semantic relationships in the dataset, giving an optimal performance. An ensemble-based deep learning approach was proposed by Aslam et al. (2021), combining BiLSTM-GRU on the LIAR dataset. The model achieved 89.8% accuracy, 91.6% recall, 91.3% precision, and 91.4% F-score, outperforming the baseline models. Contextual classification of news using CNN with BiGRU was proposed by Sarin & Kumar (2025). The model yielded an optimal result, demonstrating the robustness and quality of the ISOT dataset in news classification. Furthermore, Alghamdi et al. (2024) proposed a model that incorporated multiple features, leveraging both context and content features, and implemented BERT (for contextualised representations) and GloVe embeddings, which improved performance over the state-of-the-art model. Notable computational trade-offs exist between models; GRU-based hybrid models reduce training overhead compared to LSTM variants (Umar et al., 2023), while multilevel CNN-RNN fusion supports multimodal (text + image) classification with improved generalization (Abbas & Taeihagh, 2025).

Embedding techniques and architecture have also significantly influenced hybrid performance. Ouassil et al. (2022) combined two embedding techniques, Word2Vec (CBOW) and Word2Vec (Skip-Gram), on the WELFake dataset, with a CNN-BiLSTM architecture. The model yielded optimal performance in terms of precision and loss rate, with 97.74% accuracy. Agarwal et al. (2020) combined CNN and RNN to establish linguistic relationships through word embedding; moreover, they optimised prediction quality by tuning parameters to achieve 97.21% precision. Alghamdi et al. (2024) proposed a model that incorporated multiple features, leveraging both context and content features. BERT was used for contextualised representations, and GloVe embeddings were used for CNN and BiGRU, which yielded an improved performance over the baseline model. Dev et al.

(2024) leveraged TF-IDF and CountVectorizer for feature extraction, improving performance accuracy and convergence. Yadav et al. (2023) also utilized a 300-dimensional GloVe embedding with different deep-learning architectures, including LSTM, BiLSTM, CNN-LSTM, and CNN-BiLSTM on fake and real news datasets, as well as the allData dataset. CNN-BiLSTM achieved 97% accuracy. Hyperparameter optimization via grid search and cross-validation further mitigates overfitting, enabling CNN-LSTM to correctly predict 8953 news items for both classes, misclassifying only 27 (Utku, 2024)

Deep learning techniques have also demonstrated effectiveness in classifying language-specific datasets (Fouad et al., 2022). Showcasing adaptability across diverse linguistic characteristics. This is achieved by tailoring Natural Language Processing (NLP) models, such as Long Short-Term Memory (LSTM) networks or Transformers, to specific languages, often by leveraging pre-trained models designed for those languages. Multilingual evaluations show language and dataset-dependent suitability. CNN-LSTM detected misinformation using the Arabic AraNews dataset, achieving 91.4% accuracy, which indicates that it excels in Arabic dataset classification (Wotaifi & Dhannoon, 2023) and English dataset classification (GÜler & Gündüz, 2023). Almandouh et al. (2024) presented models based on ML, DL, and transformer on two Arabic datasets (ARABICFAKETWEETS and AFND), leveraging FastText's ability to capture subwords, context, and nuance in the Arabic dataset. BiGRU-BiLSTM achieved superior performance on both datasets across accuracy, F1-score, and loss metrics. Results show that sequential modelling was more suitable for the datasets, with RNN and its variants achieving over 90% accuracy across all metrics, compared to CNN, which had 64%. However, CNN performed better in the Korean language dataset (Lee et al., 2019), suggesting dataset size and linguistic structure as moderating factors.

Hybrid models (CNN-RNN/LSTM/BiLSTM) integrate both local and sequential features for better performance, achieving mean accuracies of 99.28% (ISOT), 60.71% (FA-KES), and 96.39% (WELFake), with CNN-BiLSTM peaking at 68% on FA-KES (Tokpa et al., 2023), an improvement on the non-hybrid models, demonstrating robustness in data-scarce contexts.

RQ3. Comparison of Hybrid and Non-Hybrid Deep Learning Models.

Tables 1–8 and Figure 2 compare model performances across three standard fake news datasets: ISOT, FA-KES, and WELFake. Table 8 aggregates the means, showing hybrids outperforming non-hybrids: +0.17% (ISOT: 99.28% vs. 99.11%), +8.6% (FA-KES: 60.71% vs. 52.11%), and +0.22% (WELFake: 96.39% vs. 96.17%), yielding a 3% mean advantage (85.46% vs. 82.46%). This highlights the superiority of hybrids, particularly on smaller datasets (Figure 2).

Table 2: Fake news detection accuracies of Non-Hybrid DL Models on the ISOT dataset

Author	Model	Accuracy (%)
(Nasir et al., 2021)	CNN	99
(Nasir et al., 2021)	RNN	98
(Güler & Gündüz, 2023)	CNN	98.02
(Saleh et al., 2021)	CNN	99.99
(Saleh et al., 2021)	RNN	99.6
(Saleh et al., 2021)	LSTM	99.78
(Chang, 2024)	BiLSTM	99
(Chang, 2024)	LSTM	98.81
(Sastrawan et al., 2022)	BiLSTM	99.95
(Sastrawan et al., 2022)	CNN	99.88
(Tokpa et al., 2023)	CNN	99
(Tokpa et al., 2023)	LSTM	98
(Tokpa et al., 2023)	BiLSTM	99
(Umar et al., 2023)	CNN	98
(Jian et al., 2024)	GRU+Attn	98.46
(Jian et al., 2024)	LSTM	98.94
(Jian et al., 2024)	BiLSTM+Attn	99.98
(Sarin & Kumar, 2025)	RNN	99.53
(Sarin & Kumar, 2025)	LSTM	99.52
(Sarin & Kumar, 2025)	GRU	99.65
(Utku, 2024)	LSTM	99.69
Mean Accuracy		99.11

Table 3: Fake news detection accuracies of Non-Hybrid DL Models on FA-KES dataset

Author	Model	Accuracy (%)
(Nasir et al., 2021)	RNN	50
(Nasir et al., 2021)	BiLSTM	50
(Umar et al., 2023)	CNN	56
(Saleh et al., 2021)	CNN	53.99
(Saleh et al., 2021)	RNN	48.57
(Saleh et al., 2021)	LSTM	47.64
(Tokpa et al., 2023)	CNN	50
(Tokpa et al., 2023)	LSTM	50
(Tokpa et al., 2023)	BiLSTM	53
(Umar et al., 2023)	GRU	58
(Alghamdi et al., 2024)	BiLSTM	50.62
(Alghamdi et al., 2024)	BiGRU(Attn)	56.79
(Alghamdi et al., 2024)	CNN	52.80
Mean Accuracy		52.11

Table 4: Fake news detection accuracies of Non-Hybrid DL Models on the WELFake dataset

Author	Model	Accuracy (%)
(Padalko et al., 2024)	BiLSTM+Attn	97.66
(Padalko et al., 2024)	BiLSTM	97.49
(Hashmi et al., 2024)	BiLSTM	98
(Hashmi et al., 2024)	LSTM	98
(Hashmi et al., 2024)	GRU	98
(Hashmi et al., 2024)	BiGRU	98
(Umar et al., 2023)	CNN	90
(Verma et al., 2021)	CNN	92
(Ali et al., 2023)	BiLSTM	98.52
(Sarin & Kumar, 2025)	RNN	93.93
(Sarin & Kumar, 2025)	LSTM	96.33
(Sarin & Kumar, 2025)	GRU	96.08
Mean Accuracy		96.17

Table 5: Fake news detection accuracies of Hybrid DL Models on the ISOT dataset

Author	Model	Accuracy (%)
(Nasir et al., 2021)	CNN-RNN	99
(Güler & Gündüz, 2023)	RNN-LSTM	99.9
(Umar et al., 2023)	CNN-GRU	99
(Umar et al., 2023)	CNN-RNN	100
(Tokpa et al., 2023)	CNN-LSTM	99
(Tokpa et al., 2023)	CNN-BiLSTM	99
(Sarin & Kumar, 2025)	CNN+BiRNN	99.65
(Sarin & Kumar, 2025)	CNN+BiLSTM	99.66
(Sarin & Kumar, 2025)	CNN+BiGRU	99.66
(Utku, 2024)	CNN-LSTM	99.91
(Abbas & Taeihagh, 2025)	CNN-RNN	97.25
Mean Accuracy		99.28

Table 6: Fake News Detection with Hybrid DL Models on FA-KES dataset

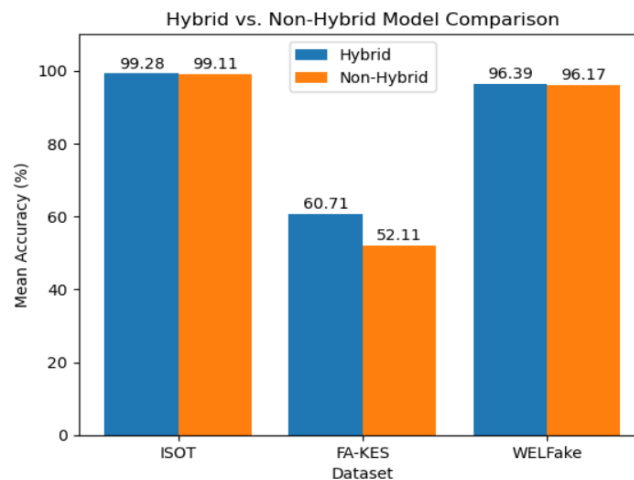
Author	Model	Accuracy (%)
(Nasir et al., 2021)	CNN-RNN	60
(Tokpa et al., 2023)	CNN-LSTM	63
(Tokpa et al., 2023)	CNN-BiLSTM	68
(Umar et al., 2023)	CNN-GRU	59
(Umar et al., 2023)	CNN-RNN	58
(Abbas & Taeihagh, 2025)	CNN-RNN	54.03
(Alghamdi et al., 2024)	CNN-BiGRU	62.96
Mean Accuracy		60.71

Table 7: Fake News Detection with Non-Hybrid DL Models on WELFake dataset

Author	Model	Accuracy (%)
(Ouassil et al., 2022)	CNN-BiLSTM	97.74
(Hashmi et al., 2024)	BiLSTM-GRU	98
(Hashmi et al., 2024)	CNN-LSTM	99
(Umar et al., 2023)	CNN-GRU	91
(Sarin & Kumar, 2025)	CNN+BiRNN	95.78
(Sarin & Kumar, 2025)	CNN+BiLSTM	95.53
(Sarin & Kumar, 2025)	CNN+BiGRU	95.86
(Abbas & Taeihagh, 2025)	CNN-RNN	98.17
Mean Accuracy		96.39

Table 8: Summary of Hybrid and Non-Hybrid Model

Dataset	Hybrid (%)	Non-Hybrid (%)
ISOT	99.28	99.11
FA-KES	60.71	52.11
WELFake	96.39	96.17
Mean Accuracy	85.46	82.46

**Fig. 2: Hybrid and Non-hybrid Model Comparison**

Discussion

This section evaluates the performance of hybrid and non-hybrid deep learning (DL) models across ISOT, FA-KES, and WELFake datasets, addressing RQ₁–RQ₃ in the result section.

Performance Analysis

Tables 1 to 8 summarize the performances of hybrid and non-hybrid deep learning models across the datasets.

Non-Hybrid Models

The high accuracies (98%-99.98%) of CNN, LSTM, and BiLSTM indicate that non-hybrid models excel on large datasets, such as ISOT, scoring a mean accuracy of 99.11%. This indicates that models that capture sequential and long-range dependencies, as well as local features, perform optimally on the ISOT dataset. This aligns with Hu et al. (2022), who note DL's strength in pattern recognition. However, the low accuracy on FA-KES suggests limitations in handling small, domain-specific datasets, possibly due to insufficient feature diversity (Nasir et al., 2021). Table 3 indicates a 52.11% mean accuracy for non-hybrid DL models. The constraints on size and specific contextual focus limit performance on the FA-KES dataset (Mridha et al., 2021). Table 4 shows that sequential models tend to be more suitable for the WELFake dataset as validated by the results of previous research (Hashmi et al., 2024; Ali et al., 2023; Padalko et al., 2024). The quality of the dataset is in its diversity and large size, with an aggregate mean accuracy of 96.17%.

Hybrid Models

Hybrid models (e.g., CNN-BiLSTM/LSTM/RNN) integrated features, capturing both syntactic patterns and long-term dependencies (Nasir et al., 2021) for a better classification, and achieved a mean accuracy of 99.28% (ISOT), 60.71% (FA-KES), and 96.39% (WELFake). CNN-BiLSTM achieved a peak accuracy of 68% on FA-KES (Tokpa et al., 2023), surpassing the performance of non-hybrid models. This robustness in data-scarce contexts can be attributed to the enhanced feature fusion. However, it comes with increased computational costs (Umar et al., 2023), posing a trade-off for real-time applications.

The aggregate accuracy across the datasets shows that the hybrid model surpassed the non-hybrid model by (+0.17%) accuracy on the ISOT dataset, (+8.6%) accuracy on the FA-KES dataset and (+0.22%) accuracy on the WELFake dataset. From Figure 2, the FA-KES dataset has lower bars and a wider difference between the hybrid and non-hybrid models bars compared to ISOT and WELFake. This indicates lower accuracy compared to ISOT and the WELFake dataset; the wider difference also suggests that hybrid models outperform non-hybrid models on the FA_KES dataset. The minimal gap between ISOT and WELFake suggests that large and diverse datasets enable feature learning for both model types. The 8.6% hybrid advantage on the FA-KES dataset reflects hybrids' ability to mitigate data scarcity, however, generalizability to low-resource languages remains untested, and higher computational demands may limit deployment in resource-constrained settings.

Limitations and Future Directions

This SLR's English-only focus and reliance on high-accuracy studies may introduce bias, limiting generality. Dataset biases and exclusion of non-peer-reviewed works further constrain the scope. Future research should explore multimodal datasets (e.g., text and images), low-resource languages, and adversarial robustness to enhance real-world applicability.

Conclusion

Deep learning techniques have demonstrated notable effectiveness in fake news detection, achieving high accuracy and efficiency by leveraging architectures such as CNNs, LSTMs, BiLSTMs, and hybrid models to enhance detection performance. This study compared hybrid and non-hybrid deep learning models for news classification. While ISOT and WELFake saw strong performance from both model types (owing to their size and diversity), the FA-KES dataset posed more challenges, yet, hybrids fared better, as they blended local feature extraction and sequence understanding, yielding better performance with accuracy gains of 0.16% to 8.6% over non-hybrids (hybrids: 60.71% aggregate mean vs. non-hybrids: 52.11%). Deployment considerations might include using hybrids for high-stakes applications and non-hybrids in resource-constrained settings. This review demonstrates that hybrid deep learning architectures outperform non-hybrid models across diverse datasets. Future research should emphasize multimodal and multilingual data, improving cross-domain robustness in fake news detection.

Conflict of Interest

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

- Abbas, F., & Taeiagh, A. (2025). A multi-level fusion-based framework for multimodal fake news classification using semantic feature extraction. *International Journal of Machine Learning and Cybernetics*, 16(9), 6531–6560. <https://doi.org/10.1007/s13042-025-02633-w>
- Abu Salem, F. K., Al Feel, R., Elbassuoni, S., Jaber, M., & Farah, M. (2019). FA-KES: A Fake News Dataset around the Syrian War. *Proceedings of the International AAAI Conference on Web and Social Media*, 13, 573–582. <https://doi.org/10.1609/icwsm.v13i01.3254>
- Adegoke, Y., Northcott, C., Leroy, A., Jari, A., & Tanko, M. (2018, November 13). *Like. Share. Kill. Nigerian police say false information on Facebook is killing people*. BBC. https://www.bbc.co.uk/news/resources/idt-sh/nigeria_fake_news
- Agarwal, A., Mittal, M., Pathak, A., & Goyal, L. M. (2020). Fake News Detection Using a Blend of Neural Networks: An Application of Deep Learning. *SN Computer Science*, 1(3), 143. <https://doi.org/10.1007/s42979-020-00165-4>
- Alarfaj, F. K., & Khan, J. A. (2023). Deep Dive into Fake News Detection: Feature-Centric Classification with Ensemble and Deep Learning Methods. *Algorithms*, 16(11), 507. <https://doi.org/10.3390/a16110507>
- Alghamdi, J., Lin, Y., & Luo, S. (2024). The Power of Context: A Novel Hybrid Context-Aware Fake News Detection Approach. *Information*, 15(3), 122. <https://doi.org/10.3390/info15030122>
- Alghamdi, J., Luo, S., & Lin, Y. (2023). A comprehensive survey on machine learning approaches for fake news detection. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-023-17470-8>
- Ali, A. A., Latif, S., Ghauri, S. A., Song, O.-Y., Abbasi, A. A., & Malik, A. J. (2023). Linguistic Features and Bi-LSTM for Identification of Fake News. *Electronics*, 12(13), 2942. <https://doi.org/10.3390/electronics12132942>

- Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Almandouh, M. E., Alrahmawy, M. F., Eisa, M., Elhoseny, M., & Tolba, A. S. (2024). Ensemble based high performance deep learning models for fake news detection. *Scientific Reports*, 14(1), 26591. <https://doi.org/10.1038/s41598-024-76286-0>
- Aslam, N., Ullah Khan, I., Alotaibi, F. S., Aldaej, L. A., & Aldubaikil, A. K. (2021). Fake Detect: A Deep Learning Ensemble Model for Fake News Detection. *Complexity*, 2021, 1–8. <https://doi.org/10.1155/2021/5557784>
- Chang, L.-J. (2024). Comparison of Machine Learning and Deep Learning Algorithms in Detecting Fake News. *28th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2024)*, 203–209. <https://doi.org/10.54808/WMSCI2024.01.203>
- Coleman, A. (2020, August 12). "Hundreds dead" because of Covid-19 misinformation. BBC Monitoring. <https://www.bbc.com/news/world-53755067>
- Deepak, S., & Chitturi, B. (2020). Deep neural approach to Fake-News identification. *Procedia Computer Science*, 167, 2236–2243. <https://doi.org/10.1016/j.procs.2020.03.276>
- Dev, D. G., Bhatnagar, V., Bhati, B. S., Gupta, M., & Nanthaamornphong, A. (2024). LSTMCNN: A hybrid machine learning model to unmask fake news. *Heliyon*, 10(3), e25244. <https://doi.org/10.1016/j.heliyon.2024.e25244>
- Fouad, K. M., Sabbeh, S. F., & Medhat, W. (2022). Arabic Fake News Detection Using Deep Learning. *Computers, Materials & Continua*, 71(2), 3647–3665. <https://doi.org/10.32604/cmc.2022.021449>
- Goldani, M. H., Safabakhsh, R., & Momtazi, S. (2021). Convolutional neural network with margin loss for fake news detection. *Information Processing & Management*, 58(1), 102418. <https://doi.org/10.1016/j.ipm.2020.102418>
- Gomede, E. (2024, January 15). *Exploring the Potential of Hybrid Deep Neural Networks*. Medium. <https://medium.com/the-modern-scientist/exploring-the-potential-of-hybrid-deep-neural-networks-8e013ea0baf6>
- Güler, G., & Gündüz, S. (2023). Deep Learning Based Fake News Detection on Social Media. *International Journal of Information Security Science*, 12(2), 1–21. <https://doi.org/10.55859/ijiss.1231423>
- Hashmi, E., Yayilgan, S. Y., Yamin, M. M., Ali, S., & Abomhara, M. (2024). Advancing Fake News Detection: Hybrid Deep Learning With FastText and Explainable AI. *IEEE Access*, 12, 44462–44480. <https://doi.org/10.1109/ACCESS.2024.3381038>
- Hu, L., Wei, S., Zhao, Q., & Liu, B. (2022). A Survey on Fake News Detection with Deep Learning. *IEEE Access*, 10, 96720–96738. <https://doi.org/10.1109/ACCESS.2022.3205472>
- Jian, W., Li, J. P., Akbar, M. A., Haq, A. U., Khan, S., Alotaibi, R. M., & Alajlan, S. A. (2024). SA-Bi-LSTM: Self Attention With Bi-Directional LSTM-Based Intelligent Model for Accurate Fake News Detection to Ensured Information Integrity on Social Media Platforms. *IEEE Access*, 12, 48436–48452. <https://doi.org/10.1109/ACCESS.2024.3382832>
- Jiang, X., Bai, X., & Yin, L. (2025). The Latest Research Progress of Attention Mechanism in Deep Learning. *Journal of Electronic Research and Application*, 9(3), 82–89. <https://doi.org/10.26689/jera.v9i3.10597>
- Kaliyar, R. K., Goswami, A., Narang, P., & Sinha, S. (2020). FNDNet – A deep convolutional neural network for fake news detection. *Cognitive Systems Research*, 61, 32–44. <https://doi.org/10.1016/j.cogsys.2019.12.005>
- Kausar, N., AliKhan, A., & Sattar, M. (2022). Towards better representation learning using hybrid deep learning model for fake news detection. *Social Network Analysis and Mining*, 12(1), 165. <https://doi.org/10.1007/s13278-022-00986-6>

- Kumar, S., Asthana, R., Upadhyay, S., Upreti, N., & Akbar, M. (2020). Fake news detection using deep learning models: A novel approach. *Transactions on Emerging Telecommunications Technologies*, 31(2). <https://doi.org/10.1002/ett.3767>
- Li, J., & Lei, M. (2022). A Brief Survey for Fake News Detection via Deep Learning Models. *Procedia Computer Science*, 214, 1339–1344. <https://doi.org/10.1016/j.procs.2022.11.314>
- Liang, H., Sun, X., Sun, Y., & Gao, Y. (2017). Text feature extraction based on deep learning: a review. *EURASIP Journal on Wireless Communications and Networking*, 2017(1), 211. <https://doi.org/10.1186/s13638-017-0993-1>
- Mahara, T., Josephine, V. L. H., Srinivasan, R., Prakash, P., Algarni, A. D., & Verma, O. P. (2023). Deep vs. Shallow: A Comparative Study of Machine Learning and Deep Learning Approaches for Fake Health News Detection. *IEEE Access*, 11, 79330–79340. <https://doi.org/10.1109/ACCESS.2023.3298441>
- Mridha, M. F., Keya, A. J., Hamid, Md. A., Monowar, M. M., & Rahman, Md. S. (2021). A Comprehensive Review on Fake News Detection With Deep Learning. *IEEE Access*, 9, 156151–156170. <https://doi.org/10.1109/ACCESS.2021.3129329>
- Nasir, J. A., Khan, O. S., & Varlamis, I. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights*, 1(1), 100007. <https://doi.org/10.1016/j.jjimei.2020.100007>
- Noh, S.-H. (2021). Analysis of Gradient Vanishing of RNNs and Performance Comparison. *Information*, 12(11), 442. <https://doi.org/10.3390/info12110442>
- Ogbuju, E., Abiodun, T., & Oladipo, F. (2023). Text Analytics Solutions for the Control of Fake News: Materials and Methods. *International Journal of Open Information Technologies*, 11(3), 69–74. <https://cyberleninka.ru/article/n/text-analytics-solutions-for-the-control-of-fake-news-materials-and-methods>
- Otor, S., Akumba, B., & Idikwu, J. (2023). A Novel Fake-News Dataset and Detection System to Mitigate Cyber War with Emphasis on Nigerian News Events. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 24–32. <https://doi.org/10.32628/CSEIT23903146>
- Ouassil, M.-A., Cherradi, B., Hamida, S., Errami, M., Gannour, O. EL, & Raihani, A. (2022). A Fake News Detection System based on Combination of Word Embedded Techniques and Hybrid Deep Learning Model. *International Journal of Advanced Computer Science and Applications*, 13(10). <https://doi.org/10.14569/IJACSA.2022.0131061>
- Padalko, H., Chomko, V., & Chumachenko, D. (2024). A novel approach to fake news classification using LSTM-based deep learning models. *Frontiers in Big Data*, 6. <https://doi.org/10.3389/fdata.2023.1320800>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. <https://doi.org/10.1136/bmj.n71>
- Ramya, S. P., & Eswari, R. (2022). Attention-Based Deep Learning Models for Detection of Fake News in Social Networks. *International Journal of Cognitive Informatics and Natural Intelligence*, 15(4), 1–25. <https://doi.org/10.4018/IJCI.295809>
- Saleh, H., Alharbi, A., & Alsamhi, S. H. (2021). OPCNN-FAKE: Optimized Convolutional Neural Network for Fake News Detection. *IEEE Access*, 9, 129471–129489. <https://doi.org/10.1109/ACCESS.2021.3112806>
- Sarin, G., & Kumar, P. (2025). Developing and examining hybrid classifiers to study social media fake news. *Social Network Analysis and Mining*, 15(1), 56. <https://doi.org/10.1007/s13278-025-01466-3>
- Sastrawan, I. K., Bayupati, I. P. A., & Arsa, D. M. S. (2022). Detection of fake news using deep learning CNN-RNN based methods. *ICT Express*, 8(3), 396–408. <https://doi.org/10.1016/j.ict.2021.10.003>

- Tokpa, F. W. R., Kamagaté, B. H., Monsan, V., & Oumtanaga, S. (2023). Fake News Detection in Social Media: Hybrid Deep Learning Approaches. *Journal of Advances in Information Technology*, 14(3), 606–615. <https://doi.org/10.12720/jait.14.3.606-615>
- Umar, Y. A., Ahmad, S. Ibrahim, & Muhammad, K. (2023). Fake News Detection Using CNN/GRU Deep Learning Model. *Sule Lamido University Journal of Science & Technology*, 7, 57–67. https://slujst.com.ng/wp-content/uploads/2023/10/SPJST368_pp-57_67.pdf
- Uniyal, M. (2024, October 18). *Feature Extraction in Machine Learning*. <https://www.applidaicourse.com/blog/feature-extraction-in-machine-learning/>
- Utku, A. (2024). Hybrid CNN-LSTM Model for Fake News Detection. *NATURENGS MTU Journal of Engineering and Natural Sciences Malatya Turgut Ozal University*. <https://doi.org/10.46572/naturengs.1571897>
- Verma, P. K., Agrawal, P., Amorim, I., & Prodan, R. (2021). WELFake: Word Embedding Over Linguistic Features for Fake News Detection. *IEEE Transactions on Computational Social Systems*, 8(4), 881–893. <https://doi.org/10.1109/TCSS.2021.3068519>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Yadav, A. K., Kumar, S., Kumar, D., Kumar, L., Kumar, K., Maurya, S. K., Kumar, M., & Yadav, D. (2023). Fake News Detection Using Hybrid Deep Learning Method. *SN Computer Science*, 4(6), 845. <https://doi.org/10.1007/s42979-023-02296-w>