International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0124
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 813 - 818

An Integrated Deep Learning Framework for Fake News and Deepfake Video Detection

B. Bharathi[1], R.Durga2, CH.Sindhupriya[3], D.Keerthana[4]
[1]Assistant Professor, Department of Computer Science and Engineering,Sri Ranganathar Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India
[2,3,4]UG - Computer Science and Engineering, Sri Ranganathar Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India
Email: *bharathi@sriet.ac.in[1], durgaravichandran384@gmail.com[2], chsindhu600@gmail.com[3], devellakeerthana11@gmail.com[4]*

Abstract
The rapid advancement of digital media technologies has led to a significant increase in the dissemination of misinformation in the form of fake news and deepfake videos. These manipulated contents pose serious threats to public trust, social stability, and information authenticity. This paper presents an integrated deep learning–based framework for detecting fake news and deepfake videos. The proposed system employs a transformer-based Natural Language Processing model to classify textual news content, while a Convolutional Neural Network–based model is used to analyze video frames for detecting facial manipulation. Additionally, a real-time detection module is implemented to analyze live video streams using face detection and temporal smoothing techniques. Experimental evaluation demonstrates that the proposed approach effectively distinguishes real and fake content with reliable accuracy. The system is designed to be modular, scalable, and suitable for real- world deployment in media verification and social networking platforms.

Keywords: *Fake News Detection,Deepfake Detection, Deep Learning, CNN, NLP, Real-Time Video Analysis*

1. Introduction
The exponential growth of online news platforms and social media has transformed the way information is consumed and shared. However, this rapid growth has also enabled the widespread dissemination of misinformation in the form of fake news and manipulated media content. Fake news can mislead the public and influence opinions, while deepfake videos exploit advanced artificial intelligence techniques to create highly realistic yet deceptive visual content. Such developments pose serious threats to public trust, social stability, and information authenticity.Traditional misinformation detection techniques rely heavily on manual verification or rule-based systems, which are inefficient when dealing with large-scale and rapidly evolving digital content. Recent advances in deep learning have shown significant potential in automatically identifying deceptive patterns in both textual and visual data. However, many existing approaches focus on either fake news detection or deepfake video detection independently, limiting their applicability in real-world scenarios.This paper proposes an integrated deep learning framework for detecting fake news and deepfake videos within a single system. The proposed approach employs a transformer- based Natural Language Processing model for classifying textual news content and a Convolutional Neural Network– based model for detecting manipulated video frames. In addition, a real-time detection module is implemented to analyze live video streams using face detection and temporal smoothing techniques.

1.1. The main contributions of this work are summarized as follows
- the design of a unified framework for both fake news and deepfake video detection,

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0124
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 813 - 818

- the application of deep learning models for text-based and video-based misinformation analysis, and
- the development of a real-time detection module suitable for practical deployment.
- The remainder of this paper is organized as follows. Section II reviews related work in fake news and deepfake detection. Section III describes the proposed system architecture and methodology. Section IV presents experimental results and analysis. Section V discusses limitations and future work, and Section VI concludes the paper.

2. Related Work

Several researchers have explored fake news detection using traditional machine learning techniques such as Naïve Bayes, Support Vector Machines, and LSTM networks. More recent studies employ transformer-based models to capture contextual semantics more effectively [1], [2].Deepfake detection approaches primarily focus on visual artifact analysis, frequency domain features, and CNN-based architectures [3], [4]. However, most existing works address either textual or video-based misinformation independently.

3. System Architecture

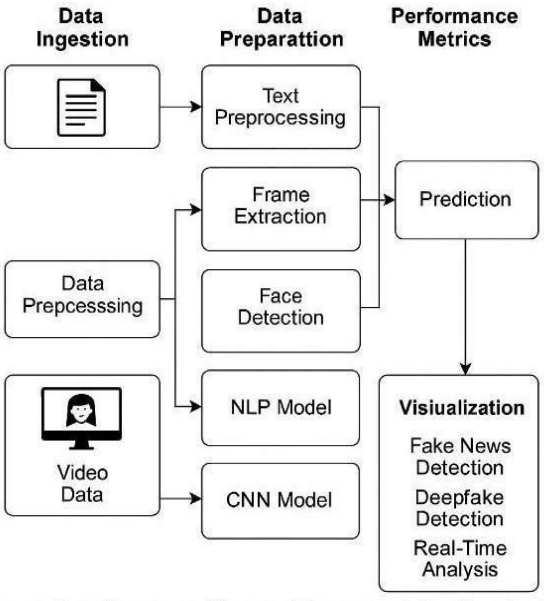

Figure 1 Overall system architecture of the proposed deep learning – based framework for fake news and deepfake video detection.

The architecture follows a modular and layered design to Shows Figure 1 Overall system architecture of the proposed deep learning – based framework for fake news and deepfake video detection. support multimodal data analysis, including textual and visual inputs.The system begins with the Data Ingestion layer, which accepts textual news content for fake news detection and video data for deepfake detection. The input data is forwarded to the Preprocessing layer, where text preprocessing operations such as tokenization and normalization are applied, while video data undergoes frame extraction and face detection to isolate relevant facial regions.The processed data is then passed to the Feature Extraction and Classification layer, which employs deep learning models for authenticity analysis. Transformer-based Natural Language Processing models are used to extract contextual semantic features from text, whereas Convolutional Neural Network (CNN) models are utilized to extract spatial features from video frames. Based on the extracted features, the system performs classification to determine whether the input content is real or fake. Finally, the results are presented through the Visualization layer, which displays the classification output along with confidence scores and additional visual information such as extracted frames and prediction graphs. This architecture enables efficient, scalable, and real-time detection of digital misinformation.

4. Proposed System

The proposed system presents an integrated deep learning– based framework for detecting fake news and deepfake videos. The system is designed to analyze both textual and visual data within a unified architecture, enabling effective identification of digital misinformation. The framework supports three modes of operation[5] fake news detection from text, deepfake video detection from uploaded videos, and real-time deepfake detection using live video streams. The overall workflow of the system includes input acquisition, preprocessing, feature extraction using deep learning models, classification, and result visualization. Each module operates independently while contributing to a unified decision-making process.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0124
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 813 - 818

4.1. Fake News Detection Module

The fake news detection module focuses on analyzing textual news content to determine its authenticity. The input text is first preprocessed through tokenization, normalization, and padding to ensure consistency. A transformer-based Natural Language Processing (NLP) model is employed to extract contextual semantic features from the text[6]. These features are then passed to a classification layer that categorizes the news content as either real or fake. The model also generates a confidence score to indicate the reliability of the prediction. This module effectively captures linguistic patterns and semantic relationships commonly associated with misinformation.

4.2. Deepfake Video Detection Module

The deepfake video detection module is responsible for detecting manipulated content in uploaded video files. Each input video is decomposed into frames at regular intervals. Facial regions are extracted from the frames using a face detection algorithm to focus the analysis on relevant visual information[7].The extracted face frames are resized and normalized before being passed to a Convolutional Neural Network(CNN). The CNN model learns spatial features such as texture inconsistencies, facial artifacts, and abnormal patterns introduced during video manipulation. Frame-level predictions are aggregated to compute an overall authenticity score, based on which the video is classified as real or fake.

4.3. Real-Time Detection Module

The real-time detection module extends the deepfake detection capability to live video streams captured through a webcam. The module continuously detects faces in each frame and processes them using the trained CNN model. To ensure prediction stability and reduce noise, temporal smoothing is applied across consecutive frames.This approach minimizes sudden fluctuations in classification results and improves reliability during real- time operation. The final prediction and confidence score are displayed to the user in real time, making the system suitable for practical deployment scenarios.

4.4. System Workflow

The overall workflow of the proposed system can be summarized as follows

- The user provides input in the form of text, video, or live camera feed.
- The input undergoes modality-specific preprocessing.
- Deep learning models extract meaningful features from the processed data.
- Classification is performed to determine authenticity.
- The results are visualized through a user-friendly interface.

4.5. Key Advantages of The Proposed System

The proposed system offers several advantages
Integrated detection of both fake news and deepfake videos. Support for real-time analysis.Modular and scalable design.Improved reliability through temporal smoothing. User-friendly visualization of results.

5. Dataset Description

The proposed system utilizes publicly available benchmark datasets to evaluate the effectiveness of fake news and deepfake video detection. Separate datasets are employed for textual and visual analysis to ensure reliable and unbiased evaluation.For fake news detection, the LIAR dataset is used. This dataset contains 12,836 short political statements collected from public sources and labeled according to their degree of truthfulness[8]. The statements are categorized into six classes: pants-fire, false, barely-true, half-true, mostly-true, and true. The dataset is pre-split into training, validation, and test sets, enabling consistent performance evaluation across different experiments.For deepfake video detection, a dataset consisting of real and manipulated facial videos is employed, inspired by the FaceForensics++ benchmark dataset. The dataset includes authentic videos as well as videos manipulated using deepfake generation techniques. These videos contain facial alterations that introduce visual artifacts and inconsistencies, which are useful for training deep learning models to distinguish between real and fake content. Video samples are processed by extracting frames at fixed intervals for frame- level analysis.The use of benchmark datasets ensures reproducibility and enables fair comparison with existing deep learning–based misinformation detection

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0124
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 813 - 818

approaches.

6. Methodology

The proposed methodology aims to detect fake news and deepfake videos using deep learning techniques. The system follows a sequential processing pipeline that includes data collection, preprocessing, feature extraction, model training, and classification. Separate methodologies are applied for textual and visual data, while the final output is presented through a unified interface.

6.1. Fake News Detection Methodology

The fake news detection process begins with the collection of textual news data from a benchmark dataset. The input text is first preprocessed to remove noise and ensure consistency. Preprocessing steps include text normalization, tokenization, and padding.After preprocessing, the text is passed to a transformer- based Natural Language Processing model, which extracts contextual and semantic features from the input. These features are then used by a classification layer to determine whether the news content is real or fake. The model outputs both the predicted class and a confidence score.

6.2. Deepfake Video Detection Methodology

For deepfake video detection, the input video is processed by extracting frames at regular intervals. Face detection is applied to each frame to isolate the facial region, which is most relevant for identifying manipulation artifacts. The extracted face images are resized and normalized before being passed to the deep learning model.A Convolutional Neural Network (CNN) is used to analyze the facial frames and extract spatial features related to texture inconsistencies and visual artifacts. Frame-level predictions are generated and aggregated to produce a final decision for the video[9].

6.3. Real-Time Detection Methodology

In real-time detection, live video frames captured from a webcam are processed continuously. Faces are detected in each frame and analyzed using the trained CNN model. To improve prediction stability and reduce noise, temporal smoothing is applied across consecutive frames. This approach ensures consistent classification during real-time operation.

6.4. Decision And Output Generation

The final classification decision is generated based on the outputs of the trained models. The system classifies the input as real or fake and computes a confidence score. The results are displayed to the user through a web-based interface, enabling easy interpretation of the system's predictions.

7. Experimental Results And Analysis

This section presents the experimental evaluation of the proposed fake news and deepfake detection framework. The performance of the system is analyzed using standard classification metrics and visualized through confusion matrix and metric comparison graphs. The evaluation focuses primarily on the deepfake video detection module at the image (frame) level, as visual artifacts are most effectively captured at this granularity.

7.1. Evalution Matrics

The performance of the proposed system is evaluated using the following standard metrics:

- **Accuracy:** Measures the overall correctness of frame-level predictions.
- **Precision:** Indicates the reliability of fake predictions by measuring how many frames predicted as fake are actually fake.
- **Recall:** Measures the ability of the system to correctly identify manipulated frames.
- **F1-Score:** Provides a balanced measure by combining precision and recall.

These metrics are widely used in image-based classification tasks and provide a comprehensive assessment of model performance.

7.2. Confusion Matrix Analysis

The confusion matrix shown in Figure 2 illustrates the frame- level classification performance of the deepfake detection model[10]. Each extracted video frame is classified as either real or fake. The matrix summarizes the number of correctly and incorrectly classified frames.

- True Positives (TP) represent fake frames correctly identified as fake.
- True Negatives (TN) represent real frames correctly identified as real.
- False Positives (FP) represent real frames incorrectly classified as fake.
- False Negatives (FN) represent fake frames incorrectly classified as real.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0124
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 813 - 818

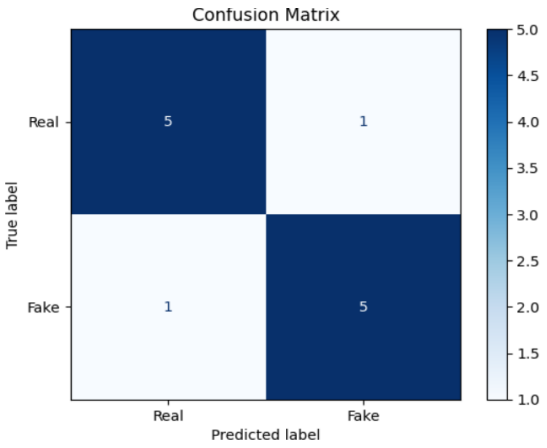

Figure 2 Confusion matrix representing frame-level classification performance of the deepfake detection model.

The results indicate that the proposed model achieves a highnumber of correct classifications, with relatively fewer misclassifications[11]. This demonstrates the effectiveness of the CNN model in identifying visual manipulation artifacts at the frame level.

7.3. Performance Metric Comparsion

To furtherr analyze the effectiveness of the proposed system, a comparison of evaluation metrics is presented in Figure 3. The bar chart illustrates the values of accuracy, precision, recall, and F1-score achieved by the deepfake detection module.

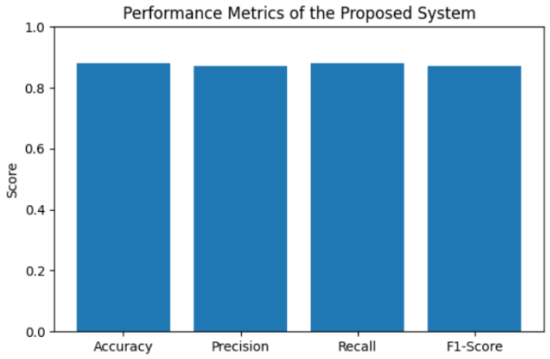

Figure 3 Comparison of accuracy, precision, recall, and F1-score for frame-level deep fake detection.

The results show that the model maintains a balanced performance across all metrics, indicating reliable detection capability. High precision reduces false alarms, while high recall ensures that most manipulated frames are successfully detected. The F1-score confirms the overall robustness of the system.

7.4. Discussion

The experimental results demonstrate that the proposed deep learning framework is capable of effectively detecting deepfake content at the image level. Frame-level evaluation allows precise analysis of visual inconsistencies introduced during manipulation. Aggregating these predictions further improves video-level decision reliability, especially in real-time scenarios.The combination of confusion matrix analysis and metric comparison provides a clear and interpretable evaluation of system performance, validating the effectiveness of the proposed approach..

Conclusion

This paper presented an integrated deep learning–based framework for detecting fake news and deepfake videos. The proposed system combines Natural Language Processing techniques for text-based fake news detection with Convolutional Neural Network–based analysis for deepfake video detection. The framework supports both offline analysis of uploaded content and real-time detection using live video streams.Experimental evaluation demonstrated that the proposed approach effectively distinguishes between real and fake content at both text and image levels. Image-based evaluation metrics, including confusion matrix, accuracy, precision, recall, and F1-score, were used to analyze frame- level performance for deepfake detection. The results indicate that the system achieves reliable classification performance and maintains stability through temporal aggregation of frame-level predictions.The modular architecture of the system enables scalability and ease of integration into real-world applications such as social media monitoring, media verification platforms, and content moderation systems. By addressing both textual and visual misinformation within a single framework, the proposed system provides a comprehensive solution to the growing challenge of digital misinformation.

Future Work

Despite its effectiveness, the proposed system can be further enhanced in several ways. Future work may

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0124
e ISSN: 2584-2854
Volume: 04 Issue: 03
March 2026
Page No: 813 - 818

include the integration of audio-based deepfake detection to enable multimodal analysis. Advanced deep learning architectures such as 3D CNNs and transformer-based video models can be explored to improve detection accuracy. Additionally, deploying the system as a cloud-based service and training it on larger and more diverse datasets may further enhance its robustness and real-world applicability.

Reference

[1]. W. Y. Wang, "Liar, liar pants on fire: A new benchmark dataset for fake news detection,"in Proc. 55th Annual Meeting of the Association for Computational Linguistics (ACL),Vancouver, Canada, 2017, pp. 422–426.

[2]. A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner,

[3]. "FaceForensics++: Learning to detect manipulated facial images,"in Proc. IEEE International Conference on Computer Vision (ICCV), 2019, pp. 1–11.

[4]. J. Thies, M. Zollhöfer, M. Stamminger, C. Theobalt, and M. Nießner,"Face2Face: Real-time face capture and reenactment of RGB videos,"in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2387–239

[5]. I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al.,"Generative adversarial nets,"in Proc. Advances in Neural Information Processing Systems (NeurIPS), 2014, pp. 2672–2680.

[6]. J. Devlin, M. Chang, K. Lee, and K. Toutanova,"BERT:Pre-training of deep bidirectional transformers for language understanding,"in Proc. NAACL-HLT, 2019, pp. 4171–4186

[7]. D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen,"MesoNet: A compact facial video forgery detection network,"in Proc. IEEE International Workshop on Information Forensics and Security (WIFS), 2018, pp. 1–7

[8]. K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu,"Fake news detection on social media: A data mining perspective,"ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.

[9]. F. Chollet,"Xception: Deep learning with depthwise separable convolutions,"in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1251–1258.

[10]. T. Wolf et al.,"Transformers:State-of-the-artnatural language processing,"in Proc. Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020.

[11]. Y. Li, M. Chang, and S. Lyu,"In Ictu oculi: Exposing AI generated fake face videos by detecting eye blinking,"in Proc. IEEE International Workshop on Information Forensics and Security (WIFS), 2018.