International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 04, April 2026)

Lightweight AI Model for Deepfake and Fake News Detection on Mobile Devices.

Sufiyan A. Ahmad[1], Syyed Musddiq[2], D. S Deshmukh[3]

[1,2]MCA II yr Sem IV P.G. Dept of Computer Applications, PRMITR Badnera, City –Amravati country-India

[3]Assistant Professor, P.G. Dept of Computer Applications, PRMITR Badnera, City –Amravati country-India

Abstract-- **Deepfake media and fake news have become major challenges in digital information systems, especially with the rapid growth of social media platforms. While existing detection methods achieve high accuracy, many of them rely on computationally expensive models that are unsuitable for mobile and resourceconstrained devices.**

This paper proposes a lightweight artificial intelligence framework for deepfake and fake news detection with a focus on efficiency and mobile deployability. The proposed system consists of two independent modules: a vision-based deepfake detection module for analyzing manipulated visual content and a text-based fake news detection module for classifying misleading textual information.

Lightweight model architectures and efficiencyoriented design principles are emphasized to reduce computational complexity and inference latency. The framework is designed to support practical deployment on mobile and edge devices while maintaining reliable detection capability. This work aims to provide a scalable and efficient foundation for future implementation and evaluation of mobile-friendly misinformation detection systems.

I. Introduction

The rapid growth of digital media and social networking platforms has significantly increased the spread of manipulated content such as deepfake videos and fake news [1], [2]. Deepfakes involve digitally altered visual or audio media that appear realistic, while fake news refers to misleading or false information presented as legitimate news [1], [2]. Both poseerious risks to public trust, social stability, and information authenticity [1].

Recent research has shown that artificial intelligence and deep learning techniques can effectively detect deepfake media and fake news [1], [2].

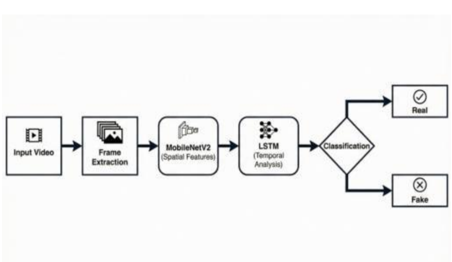

Fig 1. Overall architecture of the proposed lightweight deepfake and fake news detection system.

However, many existing approaches rely on complex and computationally intensive models that require high processing power and memory [5], [6]. Such models are often unsuitable for deployment on mobile devices and resource-constrained environments, where efficiency and low latency are critical requirements [5].

With the increasing use of mobile devices for content consumption, there is a growing need for lightweight detection frameworks that can operate efficiently without sacrificing reliability [5], [6]. Lightweight models aim to reduce computational complexity, memory usage, and inference time, making them more practical for real-world mobile and edge-based applications [6].

In this context, this paper proposes a lightweight artificial intelligence framework for deepfake and fake news detection designed specifically for mobile and resourceconstrained devices. The proposed framework integrates two modules: a vision-based module for detecting deepfake visual content and a text-based module for identifying fake news. By emphasizing efficiency-oriented design and lightweight architectures, the proposed approach aims to provide a practical foundation for future mobile-friendly misinformation detection systems.

II. Literature Review

A. Lightweight Deepfake Detection Methods

Several studies have focused on developing lightweight models for deepfake detection to reduce computational complexity while maintaining reliable performance. Surandase et al. proposed a deepfake detection approach that combines MobileNetV2 for spatial feature extraction with an LSTM network for modeling temporal information in videos [15]. The use of MobileNetV2 helps reduce model size and computation cost, making the approach suitable for resource-limited environments. However, the inclusion of temporal modeling increases inference time, which may affect real-time deployment on mobile devices.

Chen et al. introduced DefakeHop++, a lightweight deepfake detection framework designed for low-compute environments [14]. Unlike conventional deep convolutional models, this approach relies on feature-based learning to reduce model complexity while maintaining competitive detection performance.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 04, April 2026)

Experimental results on benchmark datasets indicate that the model is effective under controlled conditions, but performance degradation is observed when evaluated on unseen datasets, highlighting generalization challenges in real-world scenarios. From a mobile deployment perspective, such feature-based lightweight methods are promising, but robustness across diverse realworld data remains a critical challenge.

Al Muhaideb et al. proposed LightFakeDetect, a lightweight deepfake detection model that focuses on facial regions in video frames [17]. By limiting computation to facial areas, the approach reduces unnecessary processing of background information and improves efficiency. The model achieves competitive detection performance on standard deepfake datasets while maintaining a small model size. However, the effectiveness of the approach depends on accurate face detection and alignment, which may affect performance under challenging video conditions.

B. Fake News Detection Using Text-Based Models

Kaliyar et al. presented Fake BERT, a transformer-based approach for fake news detection in social media content [7]. The method uses contextual word representations to capture semantic relationships in text more effectively than traditional machine learning models. Experimental results reported in the study show improved classification performance on benchmark fake news datasets. However, the large model size and high computational requirements of BERT-based architectures limit their direct deployment on mobile devices.

This highlights the need for lightweight language models that balance detection accuracy with computational efficiency for mobile applications [10], [11], [12].

Singh et al. presented SEMI-FND, a multimodal fake news detection framework that combines textual and auxiliary information to improve classification performance [8]. The approach uses an ensemble-based design to capture different aspects of misinformation while maintaining faster inference compared to complex multimodal systems. Results indicate improved robustness across datasets; however, increased system complexity may impact deployment on highly resource-constrained mobile devices.

Overall, existing studies demonstrate that while deep learning-based approaches achieve strong performance in deepfake and fake news detection, achieving an effective balance between accuracy, efficiency, and mobile deployability remains an open research challenge [6]

III. Methodology

A. Overall System Architecture

The proposed system consists of two independent but complementary modules: a vision-based deepfake detection module and a text-based fake news detection module. Both modules are designed with a focus on lightweight architectures to support deployment on mobile or resourceconstrained devices [5]. The system processes visual and textual inputs separately and produces binary classification outputs indicating whether the content is real or fake.

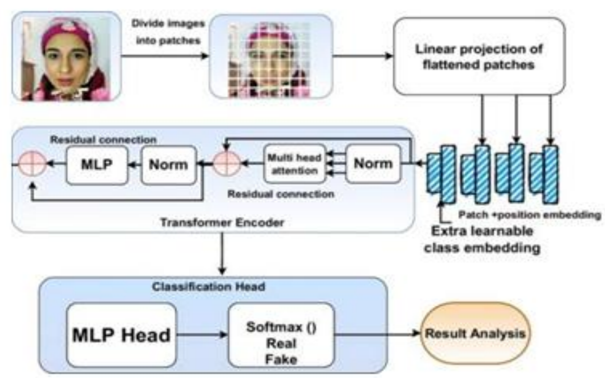

Fig 2. Overall architecture of the proposed lightweight deepfake and fake news detection system.

B. Deepfake Detection Module

The deepfake detection module operates on video inputs by extracting frames and focusing on facial regions to reduce unnecessary computation. A lightweight convolutional neural network is used for spatial feature extraction, enabling efficient processing of visual information [16], [17]. For video-based analysis, temporal consistency across frames can be modeled to improve detection reliability [15]. The module performs binary classification to distinguish between real and manipulated content.

C. Fake News Detection Module

The fake news detection module processes textual content obtained from news articles or social media posts. Text preprocessing is performed to remove noise and prepare the input for classification. A lightweight transformer-based language model is used to capture contextual information while maintaining reduced computational complexity [10], [11], [12]. The module outputs a binary label indicating whether the input text is fake or genuine.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 04, April 2026)

D. Lightweight Optimization for Mobile Devices

To ensure suitability for mobile deployment, lightweight model architectures are adopted for both vision and textbased modules. Model compression techniques such as parameter reduction and efficient architectural design are considered to minimize memory usage and inference latency [6], [23]. These optimizations enable the proposed system to balance detection performance with computational efficiency on resource-constrained devices.

E. Implementation Details

The proposed system is designed with publicly available benchmark datasets and standard evaluation practices in mind. Separate processing pipelines are considered for the deepfake detection and fake news detection modules, following commonly used experimental protocols in prior studies [3], [21], [22]. Standard evaluation metrics are intended to assess classification performance, with additional emphasis on efficiency-related factors relevant to mobile deployment.

IV. Datasets Used

Proposed system is evaluated using publicly available benchmark datasets commonly used in deepfake and fake news detection research. For deepfake detection, the FaceForensics++ dataset is used due to its wide adoption and variety of manipulated facial videos [3]. In addition, the DeepFake Detection Challenge (DFDC) dataset is considered to assess robustness under diverse and realistic conditions [21]. For fake news detection, the FakeNewsNet dataset is utilized, as it provides labeled news content along with contextual information from social media sources [22].

V. Data Preprocessing

For deepfake detection, video inputs are processed by extracting frames and detecting facial regions to reduce unnecessary background information. The extracted face images are resized and normalized before being passed to the model, following common practices in prior deepfake detection studies [16], [17]. For fake news detection, textual data undergoes basic preprocessing steps including cleaning, tokenization, and conversion into a format suitable for input to the language model [7]

The performance of the proposed system is evaluated using standard classification metrics.

Accuracy is used to measure overall correctness, while precision, recall, and F1-score are employed to provide a balanced assessment of detection performance. These metrics are widely used in existing deepfake and fake news detection studies to ensure fair and consistent comparison [1], [2].

VI. Conclusion

This paper proposed a lightweight artificial intelligence framework for deepfake and fake news detection with a focus on mobile and resource-constrained environments. The framework integrates a vision-based deepfake detection module and a text-based fake news detection module, both designed to emphasize efficiency and reduced computational complexity. By adopting lightweight model architectures, the proposed system aims to support practical deployment on mobile and edge devices while maintaining reliable detection capability.

Future work will involve implementing the proposed framework and evaluating its performance on standard benchmark datasets. Further optimization techniques such as model compression, quantization, and mobile-optimized inference frameworks can be explored to enhance efficiency and robustness. These extensions are expected to improve the applicability of the proposed framework in real-world mobile misinformation detection scenarios.

REFERENCES

[1] E. Altuncu, V. N. L. Franqueira, and S. Li, "Deepfake: Definitions, Performance Metrics and Standards, Datasets and Benchmarks, and a Meta-Review," Front. Big Data, vol. 7, Sep. 2024, doi: 10.3389/fdata.2024.1400024.

[2] X. Zhou and R. Zafarani, "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," ACM Comput. Surv., vol. 53, no. 5, Sep. 2021, doi: 10.1145/3395046.

[3] Rössler et al., "FaceForensics++: Learning to Detect Manipulated Facial Images," in Proc. IEEE/CVF ICCV, Oct. 2019, doi: 10.1109/ICCV.2019.00009.

[4] Rössler et al., "FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces," Mar. 2018, arXiv:1803.09179. doi: 10.48550/arXiv.1803.09179.

[5] R. David et al., "TensorFlow Lite Micro: Embedded Machine Learning on TinyML Systems," Mar. 2021, arXiv:2010.08678. doi: 10.48550/arXiv.2010.08678.

[6] Y. Cheng et al., "A Survey of Model Compression and Acceleration for Deep Neural Networks," Jun. 2020, arXiv:1710.09282. doi: 10.48550/arXiv.1710.09282.

[7] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," Multimed. Tools Appl., vol. 80, no. 8, pp. 11765–11788, Mar. 2021, doi: 10.1007/s11042-020-10183-2.