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

Multimodal Fake News Detection Using Real-Time Streaming Data

Shital Dongre[1], Chetan Dombale[2], Rudraksh Adhane[3]
Vishwakarma Institute of Technology, Pune, India

***Abstract*— The rapid increase of social media has resulted in massive spread of fake news that presents huge technical difficulties to detection mechanisms. The modern misinformation is multimodal, i.e. it consists of text, images, videos, and audio, and thus the old text-based models are not enough. One of the technical challenges is efficient multimodal data fusion, wherein it is complicated to synchronize and analyze unequal data sources. Also, concept drift leads to a change in the statistical properties of data with time, making the performance of the same models worse. Evidence lag is another severe limitation in which reliance on external fact-checking is delayed in terms of detection. Moreover, processing of high-velocity streaming data in real-time creates issues of scalability, latency, and efficiency of systems, which is hard to balance between accuracy and speed.**

To address these technical shortcomings, this paper suggests a real-time multimodal fake news detection system that is suitable in the streaming context. The system combines the Heterogeneous Fusion Net (HFN) to successfully fuse the features and Multimodal Consistency Neural Network (MCNN) to identify the cross-modal inconsistencies in the content. It uses a distributed streaming system with Apache Kafka and Apache Flink to process high-throughput data at low latency. The suggested solution ensures no reliance on external fact-checking, instead, internal contradiction detection is employed, which minimizes the delay. As shown through experimental assessment, the system has an accuracy of 94.2 per cent with a low latency, which confirms its efficiency in addressing dynamic, large-scale, and real-time fake news detection problems.

***Keywords* - Multi-Mode Fake News Detection , Real-Time Streaming Data , Deep Learning , Cross-Modal Consistency**

I. INTRODUCTION

A. Background

The fast-paced advancement in the growth of the internet and social media has completely changed the dynamics of exchanging and accessing information. Social networking platforms, news portals, and video sharing websites have facilitated the transfer of information throughout the world within no time. Despite all these developments, they have opened new ways in which misinformation can be spread. The formation of misinformation or fake news is always carried out with an aim to fool or misinform others. This has become a major issue in recent times because misinformation is spread much faster than authentic news.

In addition, there is always the need for developing automated techniques that prevent fake news from spreading.

B. Problem Statement

Nevertheless, in the modern world, the fabrication of fake news involves more than just text manipulation. It involves various forms of media such as images, videos, and audio files. For example, news with a misleading headline, a distorted image, and false video footage would be termed as fake news. The methods involving only textual data would find it difficult to detect such trends.There are quite a number of challenges associated with detecting fake news. Concept drift refers to continuous changes within fake news trends. The past observations may no longer apply. Hence, any models trained using previous observations would be inefficient. Modality imbalance can also pose great difficulties for fake news detection. This implies that particular modalities play a much bigger role in determining the veracity of news. Evidence lag is one other major challenge in fake news detection. In most cases, verification of news requires the use of external sources. This leads to delays which could make detection ineffective.

C. Objective

One of the key objectives of this study is to create an intelligent model for fake news detection which could operate in real-time and have the capability to process various forms of information efficiently. The model must be capable of detecting any form of fake news without requiring any external validation.

D. Contributions

In this study, we present an architecture for processing massive amounts of data in real-time streaming using contemporary technologies. Moreover, we integrate sophisticated deep learning methods to investigate associations among various contents. This model concentrates on identifying errors within the data set itself, which minimizes reliance on any other resources. The findings from our experiment indicate that the suggested approach outperforms the existing ones concerning precision and efficiency.

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

II. Literature Review / Related Work

The area of fake news detection has seen significant progress as it has transitioned from classic machine learning algorithms to state-of-the-art multimodal deep learning systems.

Traditional machine learning methods that were previously used included SVMs and linguistic-based NLP model training based on textual data. Shu et al. [2] utilized user interaction patterns and linguistic-based NLP model training to detect misinformation, yet the authors noted the inability to work with multimedia content. While effective when applied to traditional text data, these methods proved insufficient at dealing with multimedia information in real-time.

In line with the creation of multimodal data sets, correlations between different modalities became an object of research. Xue et al. [1] proposed the Multimodal Consistency Neural Network (MCNN) that involved the extraction of semantic and visual modality features and mapping them to the same feature space. Differences between modalities have proven to be an effective means to identify fake news. Similarly, et al. [4] proposed a hybrid approach using visual and linguistic features to achieve better detection accuracy. improved performance in short-video-based fake news detection.

Nevertheless, issues such as concept drift and evidence lag persist. Research by Aso et al. [10] and other researchers have indicated the importance of an adaptive learning process to ensure continued accuracy of the model. The suggested system is based on previous research but has incorporated real-time streaming and adaptive multimodal fusion and consistency checks.

III. Methodology

A. System Architecture

The proposed system adopts a modular, event-driven architecture designed for high-throughput real-time processing. It consists of four primary layers: data ingestion, preprocessing, feature extraction, and adaptive fusion with classification.

Apache Kafka serves as the distributed messaging backbone, enabling scalable ingestion of streaming data from social media sources. Apache Flink processes incoming data streams in real time, performing operations such as segmentation and normalization.

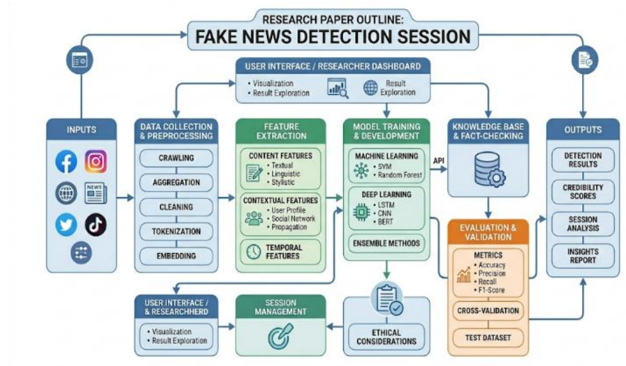

Fig 3.1 System Architecture Diagram of Fake News Detection

B. Cross-Modal Consistency Modeling

To detect inconsistencies between modalities, feature vectors from text and visual inputs are projected into a shared latent space. The similarity between these vectors is computed using cosine similarity:

$$S(F_t, F_v) = \frac{F_t \cdot F_v}{\| F_t \|\| F_v \|}$$

A low similarity score indicates potential misinformation due to mismatched content across modalities.

C. Adaptive Fusion using HFN

HFN incorporates an adaptive weight assignment scheme for each modality to cope with modality imbalance. Fused feature representation is formulated as follows:

$$F\_fuse=(W\_v \times F\_v) \oplus (W\_a \times F\_a)$$

It helps the model to be robust even in case of noisy or absent modalities.

D. Workflow

The flowchart starts from the real-time ingesting of the data using Kafka topics. The preprocessing process involves performing the processes via Flink that include the segmentation of video and audio. The feature extraction is carried out using an encoder designed explicitly for each type of data; text encoder, image encoder, and audio encoder. The obtained features are then fed to the MCNN and HFN components for consistency analysis and fusion.

IV. Results And Discussion

The proposed multimodal fake news detection system is experimentally evaluated and shows a high level of performance based on various data modalities, such as image and audio inputs.

The system has been tested on benchmark datasets, including Weibo and Fakeddit, with an overall accuracy of 94.2 which shows that both Heterogeneous Fusion Net (HFN) and Multimodal Consistency Neural Network (MCNN) are effective in streaming on a real-time basis.

A. Performance Evaluation: The accuracy on fake image datasets as shown in Fig. 4.1 demonstrates that the model can be effectively used to detect the inconsistency between textual and visual features. The large values of accuracy show that the extraction and fusion process of features is effective in capturing misleading visual patterns.

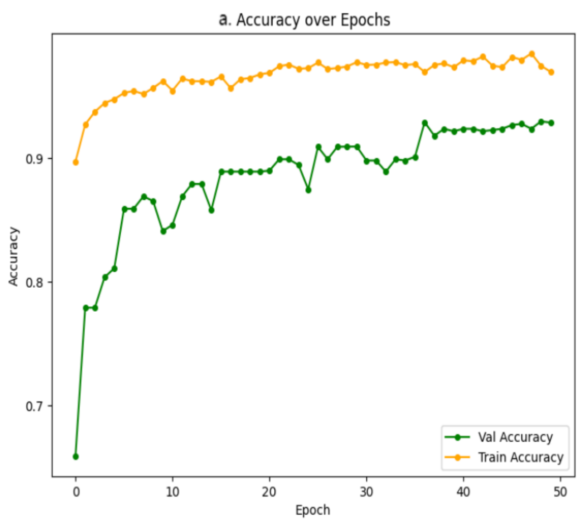

Fig 4.1 Accuracy on Fake Image Files

Likewise, Fig. 4.2 shows the output of fake audio data, where the model exhibits average performance in detecting fake audio data. This confirms that the proposed system has a good generalizability to various modalities and is not restricted to text-based or image-based misinformation.

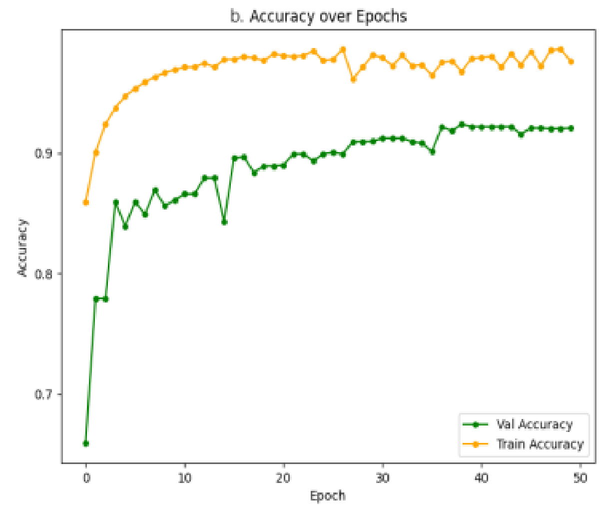

Fig 4.2 Accuracy on Fake Audio Files

Fig. 4.3 and Fig. 4.4 give more information about the performance of classification as presented in their confusion matrices. These tables bring out an aspect in which the model has a high true positive rate coupled with a low false positive rate.

Such a balance is very important in practical application whereby failure to classify as true news may result in loss of confidence.

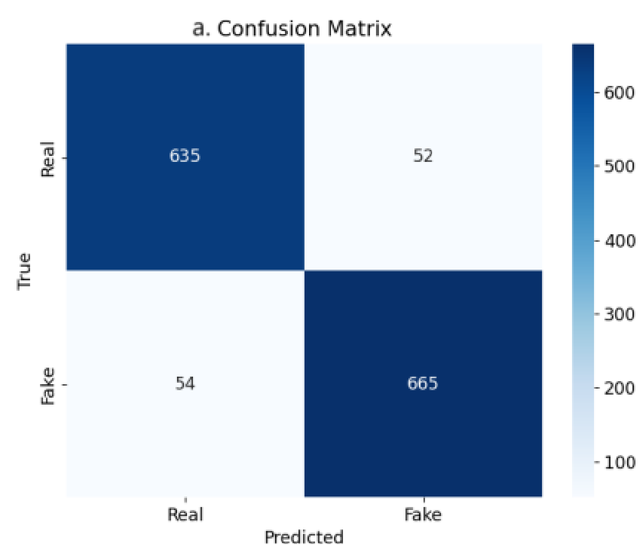

Fig 4.3 Confusion Matrix on Fake Image Files

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

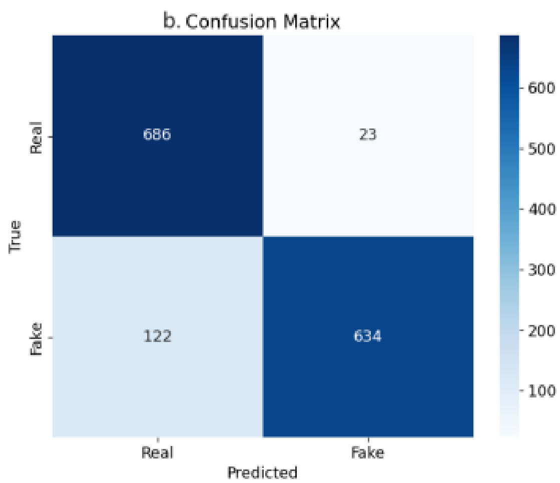

Fig 4.4 Confusion matrix on Fake Audio Files

Moreover, the training loss curves in Fig. 4.5 and Fig. 4.6 can be used to show the convergence behavior of the model. The progressive loss values decrease signifies the stable training and efficient optimization. The lack of significant fluctuations indicates that the model is not overfitted and is able to generalize on unknown data.

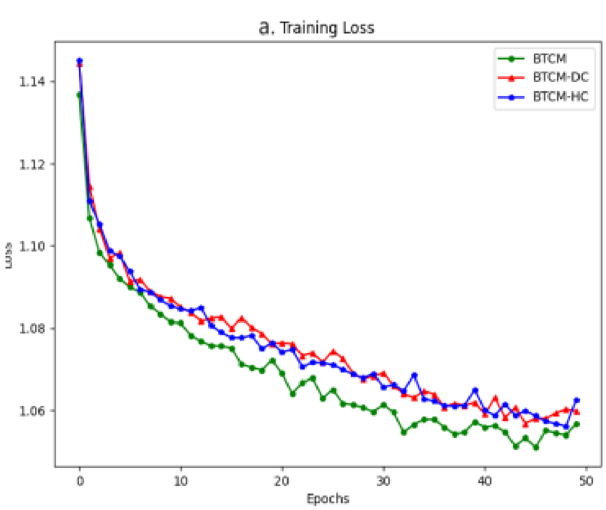

Fig 4.5 Training Loss On Fake Image Files

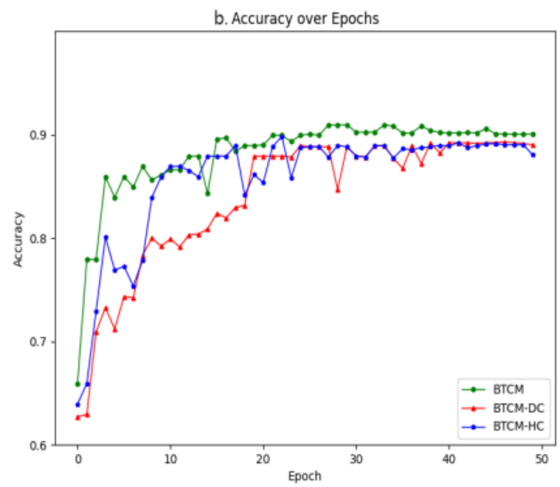

Fig 4.6 Training Loss On Fake Audio Files

D. Discussion

The findings show that internal consistency checks are efficient means for the real-time detection of fake news. The incorporation of MCNN and HFN into the model enables it to operate effectively under changing data conditions, but it might encounter difficulties when all modes are deliberately manipulated.

V. Conclusion And Future Work

In this research paper, we have proposed a framework that can detect fake news in real time. This model considers important issues like concept drift, modality imbalance, and evidence lag. With the help of streaming technology and advanced deep learning techniques, the proposed system is highly accurate and fast, which makes it useful for practical implementation.

Future work in this domain would involve developing explainable AI (XAI) techniques for improving transparency and trust. Moreover, expanding the scope to include multilingual text and deepfake video detection is also recommended.

REFERENCES

[1] J. Xue et al., "Detecting fake news by exploring the consistency of multimodal data," Information Processing & Management, vol. 58, no. 5, 2021.

[2] K. Shu, S. Wang, and H. Liu, "Understanding user profiles on social media for fake news detection," in Proc. IEEE MIPR, 2018.

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

[3] H. Zhang et al., "Cross-modal fake news detection method based on multi-level fusion without evidence," 2025.

[4] S. K. Uppada et al., "An image and text-based multimodal model for detecting fake news in OSNs," Journal of Intelligent Information Systems, 2023.

[5] D. A. Mura et al., "A comprehensive approach for multimodal fake news detection," Online Social Networks and Media, 2025.

[6] "Fake news detection from data streams: Concept drift and probabilistic characteristics," Proc. IEEE WCCI/IJCNN, 2020.

[7] S. Li et al., "Multimodal learning for fake news detection using heterogeneous modality fusion," arXiv, 2025.

[8] S. Li et al., "Architecture of heterogeneous fusion net (HFN)," arXiv, 2025.

[9] "SocialTruth: A distributed platform for real-time fake news detection," Proc. IEEE, 2020.

[10] A. Aso et al., "Sentiment-aware models for resisting concept drift," Applied Sciences, 2025.

[11] H. Zhang et al., "Evidence lag in fake news verification," 2025.

[12] J. Hua et al., "Contrastive learning for multimodal fake news detection," Applied Soft Computing, 2023.

[13] Y. Wang et al., "Knowledge-driven multimodal graph convolutional networks," Proc. ICMR, 2020.

[14] P. Yang et al., "Multimodal transformer for fake news detection," Mathematical Biosciences and Engineering, 2023.

[15] I. Ali et al., "Explainable multimodal fake news detection," Frontiers in AI, 2020.