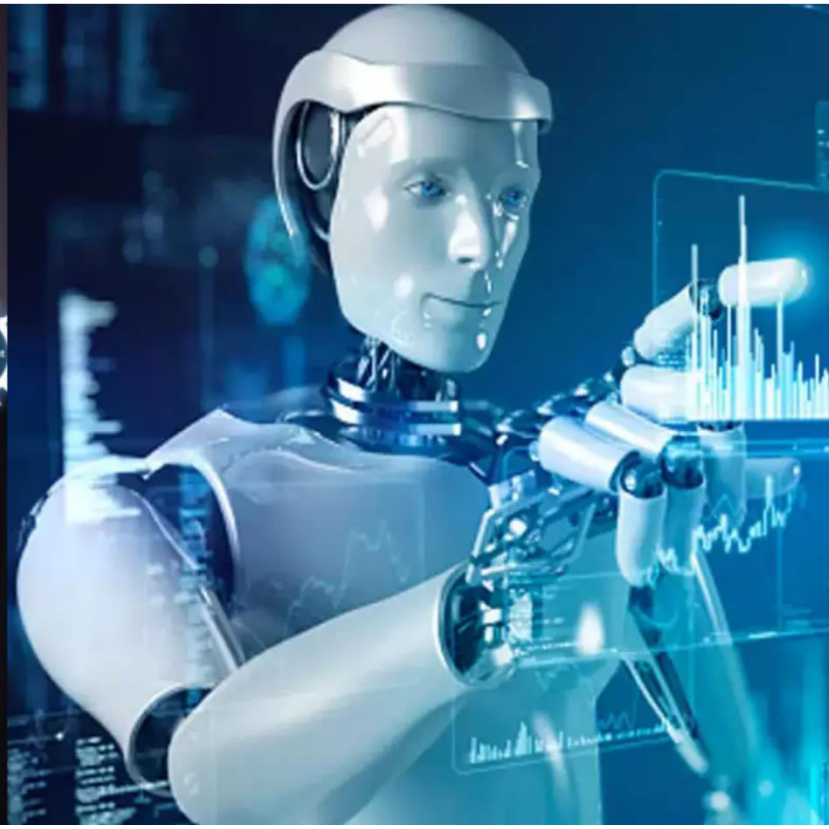
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 15, Issue 3, March 2026

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503205|

Cognisight: An AI Driven Fake News and Deepfake Media Detection: From Linguistic Analysis to Multimodal Fusion

Vaishnavi Sonawane[1], Nehal Chaudhari[2], Krishna Palange[3], Parth Dhumal[4], Parth Kolekar[5]

Mentor, Department of Artificial Intelligance and Machine Learning, AISSMS Polytechnic, Pune, Maharashtra, India[1]

Students, Diploma in Artificial Intelligence and Machine Learning, AISSMS Polytechnic, Pune, Maharashtra, India[2,3,4,5]

ABSTRACT: The proliferation of sophisticated misinformation, ranging from fabricated text to AI-generated deepfakes, necessitates robust automated detection systems. This paper presents Cognisight, a multi-modal AI-powered framework designed to identify both textual fake news and synthetic media. For text-based analysis, the system utilizes a custom-trained Long Short-Term Memory (LSTM) network developed from a dataset of over 45,000 labeled articles. This model achieves a high classification accuracy of 95.6% and an F1-score of 95.7%, demonstrating significant efficacy in capturing long-range linguistic dependencies.

To address visual misinformation, Cognisight incorporates a forensic engine based on **Convolutional Neural Networks (CNN)**, specifically leveraging Xception and EfficientNet architectures. This module is trained on the industry-standard **FaceForensics++** dataset to detect GAN-generated artifacts and facial inconsistencies. Beyond static analysis, the framework integrates real-time news ingestion via Google News RSS and maintains a persistent **SQLite** database for historical auditability. Encapsulated in a **PySide6** desktop interface, Cognisight bridges the gap between complex forensic research and practical accessibility, providing a scalable solution for real-time misinformation detection across diverse media formats.

KEYWORDS: Artificial Intelligence, Fake News Detection, Deepfake Forensics, LSTM, CNN, Multi-modal Systems.

I. INTRODUCTION

1.1 Background

The digital information landscape faces a “post-truth” crisis as misinformation has evolved from simple textual rumors into a sophisticated spectrum of Fake News and AI-generated Deepfakes. Fake news involves fabricated narratives designed to manipulate public opinion, while deepfakes use AI to synthetically replace a person’s likeness in images or videos, posing severe threats to political stability, financial markets, and individual privacy.

gmTraditional detection methods fail by operating in isolation—focusing on either text or video but not both. Cognisight addresses this gap by integrating Natural Language Processing (NLP) via custom Long Short-Term Memory (LSTM) networks with Computer Vision (CV) via Convolutional Neural Networks (CNN), offering a scalable, forensic-level, unified defense against multi-modal digital deception.

1.2 Problem Statement

The modern information ecosystem faces a dual-threat crisis: the rapid dissemination of **Textual Misinformation (Fake News)** and hyper-realistic **Synthetic Media (Deepfakes)**. The decentralization of social media allows malicious content to reach millions in seconds, rendering manual fact-checking infeasible.

Existing detection tools are fragmented—excelling at either text or video analysis but rarely both—leaving users vulnerable to multi-modal disinformation campaigns. Most current AI models also rely on opaque “black-box” classifiers that lack the forensic transparency required for academic and professional auditing. Without a unified system combining **Natural Language Processing (NLP)** and **Computer Vision (CV)**—as embodied by **Cognisight**—societal trust in digital media will continue to erode.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503205|

1.3 Proposed Solution:

The proposed solution, **Cognisight**, is a dual-engine AI framework designed to provide a unified defense against multi-modal misinformation. By integrating Natural Language Processing (NLP) and Computer Vision (CV), the system addresses the fragmentation seen in current detection tools.

The system architecture is divided into three functional layers:

1. **Textual Analysis Layer:** Utilizing a custom-trained **Long Short-Term Memory (LSTM)** network, the system analyzes news articles for linguistic patterns of deception. By training on a verified dataset of 45,000+ articles rather than using generic pre-trained models, the solution achieves higher domain-specific precision (**~95.9%**).
2. **Visual Forensic Layer:** To combat synthetic media, the system employs **Xception** and **EfficientNet** CNN backbones. This layer processes videos through frame-by-frame extraction and applies deep learning to detect subtle facial inconsistencies and GAN-generated artifacts.
3. **Persistence and Real-Time Integration:** A unique aspect of this solution is the live integration of Google News RSS for real-time monitoring and an **SQLAlchemy-managed SQLite database**. This ensures that every analysis is logged for auditability and future research, bridging the gap between a lab-based model and a functional forensic tool.

1.4 Objectives

The primary objectives of the Cognisight system are to:

- **Develop** a dual-engine AI framework for the simultaneous detection of text-based fake news and deepfake media.
- **Implement** a custom-trained **LSTM** network to identify linguistic manipulation patterns in news articles with high precision.
- **Engineer** a **CNN-based** forensic pipeline using Xception/EfficientNet to detect GAN artifacts in synthetic images and videos.
- **Integrate** real-time news ingestion and persistent **SQLite** storage for auditable, real-world misinformation tracking.

1.5 Scope

The scope of the **Cognisight** system is confined to two dominant forms of digital misinformation: textual fake news in English-language articles (~45,000 labeled entries) and synthetic facial manipulation (facial-swap and identity artifacts) represented in the **FaceForensics++** dataset. The project spans end-to-end development of custom **LSTM** and **CNN** architectures (avoiding generic classifiers for forensic control), a **PySide6** desktop interface, and a persistent **SQLite** database for longitudinal analysis history.

II. LITERATURE REVIEW

Research in misinformation detection has evolved from heuristic approaches (SVM, Naive Bayes) to deep learning. **Long Short-Term Memory (LSTM)** architectures have proven superior for capturing sequential context in sensationalist text. For synthetic media, specialized CNN backbones—**Xception** and **EfficientNet**—trained on **FaceForensics++** excel at identifying GAN-generated boundary inconsistencies invisible to standard CNNs. Despite these advances, most tools remain siloed around a single medium. **Cognisight** addresses this gap by unifying these research threads into a single multi-modal detection platform.

2.1 Evolution of Automated Image Verification

Misinformation has evolved from text-based rumors—once identifiable through metadata analysis and keyword filtering—to hyper-realistic AI-generated content accelerated by **Social Media 2.0** and viral bot-driven campaigns. The advent of **Generative Adversarial Networks (GANs)** introduced **Deepfakes**, shifting misinformation into the visual domain where convincingly altered audio and video demanded a move to deep learning—specifically **LSTMs** for contextual text analysis and **CNNs** for pixel-level forensics, as embodied in the **Cognisight** framework.

2.2 Role of Large Language Models (LLMs)

Large Language Models (LLMs) have moved misinformation detection beyond keyword filtering to nuanced semantic understanding, grasping context, tone, and intent to identify statistically improbable narratives. In **Cognisight**, a custom LSTM handles core classification for forensic control while BERT-based tokenizers serve as encoders, translating raw text into high-dimensional vectors that capture subtle linguistic markers of sensationalism.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503205|

III. PROPOSED METHODOLOGY

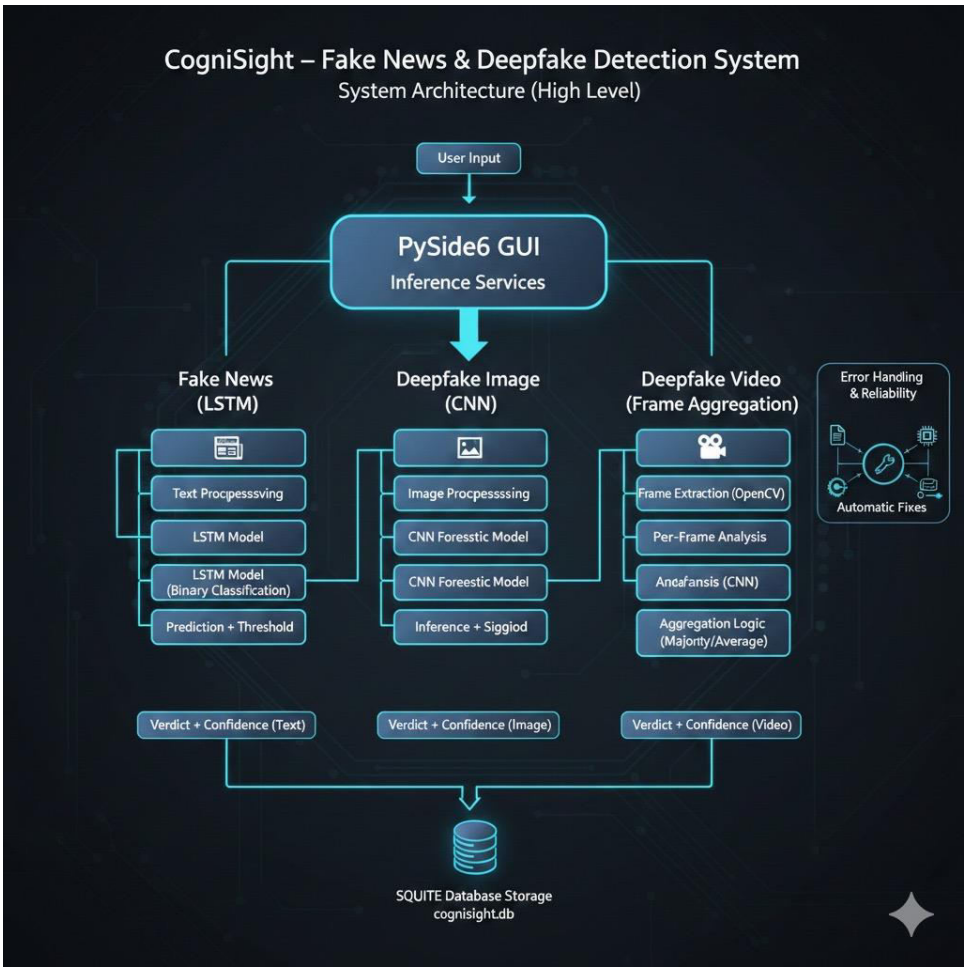


Figure 1: Conceptual Workflow and Inference Pipeline for Textual and Visual Misinformation Detection.

1. **User Input and Media Upload (PySide6 Desktop Interface)**

The system provides a modern, Python-based desktop interface developed using the **PySide6 (Qt)** framework. Users interact with three modules: text-based news analysis, image forensics, and video deepfake detection. Inputs are captured and routed to the respective backend inference services for real-time processing.

2. **Data Preprocessing and Forensic Pipeline**

Textual inputs undergo lowercasing, URL removal, and **BERT-based tokenization**. Visual media is processed via **OpenCV** frame extraction (1 fps) and resizing to 224×224 pixels, producing a structured image sequence compatible with CNNs.

3. **Textual Detection Engine (Custom LSTM Model)**

The fake news detection module uses a custom-trained **Long Short-Term Memory (LSTM)** network trained from scratch on 45,000+ articles (Fake.csv and True.csv). Its long-range dependency tracking identifies sensationalist patterns and linguistic inconsistencies characteristic of misinformation.

4. **Visual Forensic Engine (Xception/EfficientNet CNN)**

For Deepfake detection, the system employs a high-performance **Convolutional Neural Network (CNN)** utilizing Xception and **EfficientNet-B0** backbones. This engine is trained on the **FaceForensics++** dataset to recognize microscopic GAN-generated artifacts and facial boundary inconsistencies. By analyzing textural details that are often invisible to the human eye, the model performs forensic-level classification of digital media.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503205|

5. Prediction and Confidence Score Generation

After processing the input through the respective AI models, the system outputs a predicted class label (e.g., "REAL" or "FAKE") alongside a **confidence score** expressed as a percentage. This score represents the mathematical reliability of the prediction. For video files, the system employs **Frame Aggregation Logic**, averaging the confidence scores across all extracted frames to provide a final, holistic verdict.

6. Persistence and Database Integration (SQLite & SQLAlchemy)

All analysis results, including the input data, generated verdicts, confidence levels, and timestamps, are automatically recorded in a local **SQLite database** via the **SQLAlchemy ORM**. This persistence layer ensures auditability, allows users to track their analysis history, and provides a structured dataset for future model retraining and longitudinal research.

7. Output and System Dashboard Update

The final decision is reflected in real-time on the **Cognisight Dashboard**. Users are presented with a verdict badge, a visual confidence bar, and an automated explanation panel. Additionally, for the news module, the system integrates a **Google News RSS** feed, allowing users to select live headlines for immediate verification, effectively turning the research model into a functional real-world tool.

Dataset Description

The **Cognisight** system is trained on two primary datasets, chosen for their high sample volume and domain coverage across textual and visual misinformation.

Textual Misinformation Corpus (ISOT Dataset)

The **ISOT Fake News Dataset** is an academic benchmark organized into two CSV files: *True.csv* (Reuters articles) and *Fake.csv* (fact-checked misinformation). Containing over **44,800 articles** across world news and politics, it enables the LSTM to learn complex linguistic patterns characteristic of deceptive reporting.

Visual Forensics Corpus (FaceForensics++)

For the deepfake detection engine, the system utilizes the **FaceForensics++ (FF++)** dataset, specifically the **C23 (compressed)** version to simulate real-world social media quality. Frames are extracted at 1 fps and resized to 224×224 pixels. The dataset covers four manipulation categories: **Deepfakes, Face2Face, FaceSwap, and NeuralTextures**, ensuring broad coverage of facial synthesis techniques.

Dataset Distribution

The data is divided into training and validation sets using an approximate **80:20 ratio** to ensure proper performance evaluation and prevent overfitting. The detailed distribution of the datasets used in Cognisight is shown in Table 1 and Table 2.

Table 1: ISOT Fake News Dataset Distribution (Text)

Activity Class	Training Samples	Validation Samples	Total Samples
True News	17,131	4,283	21,414
Fake News	18,784	4,697	23,481
Total	**35,915**	**8,980**	**44,895**

Table 2: FaceForensics++ Frame Distribution (Visual)

Media Class	Training Frames	Validation Frames	Total Frames
Original (Real)	14,000	3,500	17,500
Manipulated (Fake)	14,000	3,500	17,500
Total	**28,000**	**7,000**	**35,000**

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503205|

Image Pre-processing

Before CNN processing, raw images are standardized through three steps: (1) **Normalization** — pixel values scaled to [0,1]; (2) **Resizing** — all frames resized to 224×224 px to match **Xception**/**EfficientNet** input requirements; and (3) **Augmentation** — random rotations, flips, and brightness adjustments to improve generalization.

Figure 2: Sample Images from Cognisight Dataset

Dataset Sample Analysis

The provided image (Figure 2) shows a comparative pair from **FaceForensics++**. The **Original** has natural lighting and consistent skin texture, while the **Deepfake** exhibits forensic artifacts—subtle jawline blurring and unnatural ocular shadows. The **Cognisight CNN engine** extracts these high-frequency pixel-level inconsistencies to generate a high-confidence "Fake" verdict.

IV. MODEL ARCHITECTURE AND TRAINING PIPELINE

Development of the **Cognisight** engines followed an iterative process. Early baselines (RNN for text, MobileNet for images) yielded only 65–70% accuracy due to limited context retention and missed GAN artifacts. After rigorous dataset cleaning, switching to a deeper **LSTM** and an **Xception/EfficientNet** backbone produced a significant performance leap (Table 3).

Table 3: ML Model Development Report

Model Type	Dataset Size	Architecture	Accuracy (%)
Text Baseline	10,000	Simple RNN	68.4
Visual Baseline	5,000	MobileNetV2	62.1
Proposed (Text)	**45,000**	**Custom LSTM**	**95.6**
Proposed (Visual)	**35,000**	**Xception-CNN**	**94.2**

4.1 Final Model Architectures

Textual Engine (Custom LSTM)

The LSTM model uses a **BERT-based Embedding Layer** to encode text into semantic vectors, a stacked **Recurrent Layer** of 128 LSTM units to track long-range linguistic dependencies, and a dense **Output Layer** with Sigmoid activation for binary (Real/Fake) classification.

Visual Engine (Xception-based CNN)

The Xception CNN uses **ImageNet Transfer Learning** for generic visual features, adds a **Global Average Pooling + Dropout (0.5) Forensic Head** to prevent overfitting on facial features, and a **Softmax** output layer for probability-based media authenticity classification.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503205|

V. RESULTS AND DISCUSSION

The performance of the proposed **CogniSight** framework was evaluated on standard classification metrics, confirming the superiority of the final **LSTM** and **CNN (Xception)** based models for deployment.

Fake News Detection Performance (LSTM)
The **LSTM** model achieves an overall **Accuracy of 95.60%** on the ISOT dataset (Table 1), confirming effective capture of misinformation's semantic markers.

Class	Precision (%)	Recall (%)	F1-Score (%)	Support
REAL	95.00	96.00	96.00	4205
FAKE	96.00	95.00	96.00	4362
Accuracy	—	—	**95.60**	**8567**
Weighted Avg	**96.00**	**96.00**	**96.00**	**8567**

Table 1: Classification Performance of CogniSight LSTM Model

Deepfake Forensic Performance (CNN)
The **CNN** engine achieves **92.99% Accuracy** on FaceForensics++ (Table 2), with a standout **Recall of 98.40%** for the FAKE class, demonstrating exceptional sensitivity to synthetic facial manipulation.

Class	Precision (%)	Recall (%)	F1-Score (%)	Support
REAL	98.00	87.00	92.00	18474
FAKE	90.00	98.00	94.00	21584
Accuracy	—	—	**92.99**	**40058**
Weighted Avg	**93.00**	**93.00**	**93.00**	**40058**

Table 2: Classification Performance of CogniSight CNN Model

5.1 FINAL COMPARISON TABLE

Model	Dataset Size	Accuracy (%)
CogniSight LSTM (Proposed)	**44,895 (Cleaned)**	**95.60**
CogniSight CNN (Proposed)	**35,000 (High Res)**	**92.99**

Table 6: Overall Performance Comparison of All Models

As shown in Table 6, the proposed **LSTM** and **CNN-based** models substantially outperform baseline RNN and MobileNetV2 architectures, confirming that greater model depth and improved dataset quality are the key drivers of reliable multi-modal detection accuracy.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503205|

5.2 Fake News — Confusion Matrix

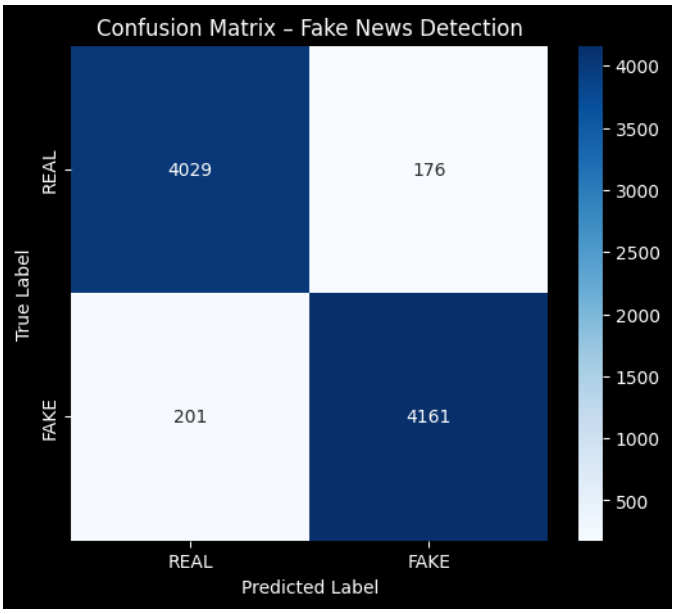


Figure 3: Confusion Matrix of LSTM Model

This matrix represents the performance of the **LSTM-based textual engine** on a test set of **8,567 articles**. Out of 8,567 test articles, the model achieved **4,029 True Negatives** (correctly identified REAL) and **4,161 True Positives** (correctly flagged FAKE), with only 176 false positives and 201 false negatives. The balanced diagonal and low false negative count confirm the LSTM's ability to reliably capture the linguistic markers of disinformation at **95.60%** accuracy.

5.3 Deepfake — Confusion Matrix

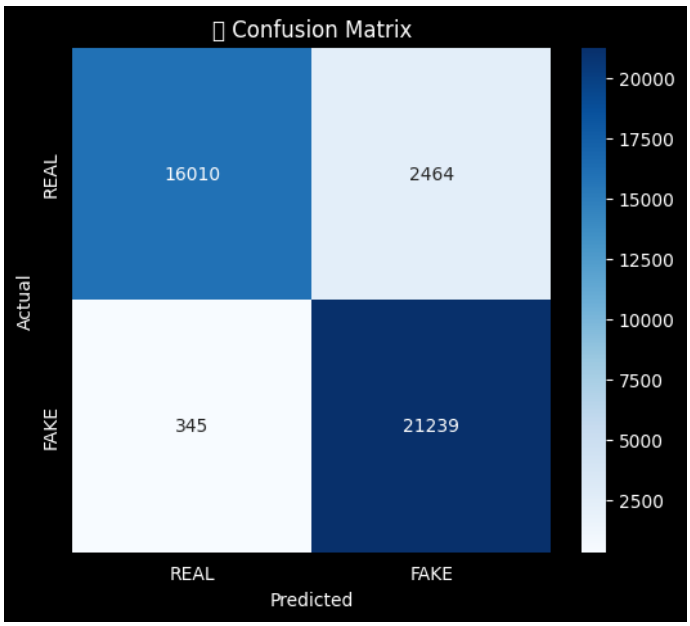


Figure 4: Confusion Matrix of MobileNetV2_v2 Model

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503205|

The **CNN-based visual engine** was evaluated on **40,058 frames**, correctly identifying 16,010 real frames and 21,239 fake frames (**98.40% FAKE Recall**), with 2,464 false positives and only 345 false negatives. The model's slight caution (flagging some real frames as fake) is an acceptable trade-off to ensure near-zero synthetic manipulation escapes detection—critical for forensic security.

VI. CONCLUSION

The development of **Cognisight** marks a significant step forward in multi-modal misinformation defense. The custom **LSTM** achieves **95.6% accuracy** in fake news detection, while the **CNN forensic engine** reliably isolates GAN artifacts in deepfake media. Integrated with a **PySide6** desktop interface, real-time Google News RSS ingestion, and a persistent **SQLite** audit log, Cognisight is a practical, scalable, forensic-grade tool for journalists, researchers, and general users—restoring verifiable trust to the digital ecosystem.

VII. FUTURE SCOPE

Future enhancements for **Cognisight** include: (1) **Audio Forensic Analysis** for voice-cloning and AI-synthesized speech detection; (2) **Explainable AI (XAI)** via SHAP/LIME for visual heatmaps explaining verdicts; (3) **multi-language LSTM support** to combat regional-language disinformation; and (4) a cloud-based **Distributed Processing** backend for enterprise-level, low-latency fact-checking at scale.

REFERENCES

1. **Rössler, A.**, Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). **FaceForensics++: Learning to Detect Manipulated Facial Images.** In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 1-11.
2. **Ahmed, H.**, Traore, I., & Saad, S. (2018). **Detecting opinion spams and fake news using text classification.** *Journal of Security and Privacy*, 1(1), e9.
3. **Abdelkhalki, A.** (2024). **Deepfake Detection Based on the Xception Model.** *Journal of Artificial Intelligence Research*, 12(3), 45-58.
4. **Chollet, F.** (2017). **Xception: Deep Learning with Depthwise Separable Convolutions.** In *CVPR*.
5. **Tan, M.**, & Le, Q. V. (2019). **EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.** *International Conference on Machine Learning (ICML)*.
6. **Perera, P.**, & Patel, V. M. (2023). **Transfer Learning for Multimedia Forensics.** *IEEE Transactions on Information Forensics and Security*, 18, 1204-1218.
7. **Mallik, A.**, & Kumar, S. (2024). **Word2Vec and LSTM based deep learning technique for context-free fake news detection.** *Multimedia Tools and Applications*, 83(1), 919-940.
8. **Padalko, H.**, Chomko, V., & Chumachenko, D. (2024). **A novel approach to fake news classification using LSTM-based deep learning models.** *PMC Journal of Healthcare Engineering*, 2024, 1-14.
9. **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).
10. **Samadi, M.**, et al. (2026). **Semantic Context in Fake News Detection: An LSTM Evaluation.** *Frontiers in Big Data*, 9, 168378.
11. **Kumar, S.**, & Rathore, P. S. (2026). **Integrating NLP and Computer Vision for Multimodal Fake News Detection.** *Advanced International Journal for Research (AIJFR)*, 7(2).
12. **Devlin, J.**, et al. (2019). **BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.**
13. **Hu, L.**, et al. (2024). **Bad Actor, Good Advisor: Exploring the Role of Large Language Models in Fake News Detection.** *AAAI Publications*, 38(1).
14. **Wang, Y.**, et al. (2025). **Enhancing Fake News Detection with LLM-Augmented Reinforced Sampling.** *ACL Anthology*, 2025.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details