

# AI Powered Digital Forensics & Crime Investigation Supporting System

PAVITHRA NATARAJAN<sup>1</sup>, SRIDHARAN TR<sup>2</sup>, MOHAMED IBRAHIM B<sup>3</sup>, SHIVA PRAKHASH K<sup>4</sup>

<sup>1</sup>Assistant Professor, Department of Artificial Intelligence & Data Science, Science Dhanalakshmi Srinivasan University, Trichy, India

<sup>2,3,4</sup>Department of Artificial Intelligence & Data Dhanalakshmi Srinivasan University, Trichy, India

**Abstract-** The increasing complexity of modern crime investigation demands intelligent systems capable of analyzing diverse data sources efficiently. This paper presents a Crime AI Support System, a multi-modal framework integrating facial emotion recognition, voice analysis, media tampering detection, and an AI-based chatbot. The system leverages computer vision, speech processing, and digital forensic techniques to provide real-time insights into emotional states and evidence authenticity. A Flask-based backend enables modular integration and efficient processing of multimedia inputs. The chatbot offers contextual assistance using a hybrid rule-based and AI-driven approach. Experimental results indicate improved accuracy in emotion detection and reliable identification of tampered media. The proposed system enhances investigative efficiency and provides effective support for victims through integrated artificial intelligence solutions.

**Index Terms-** Artificial Intelligence, Crime Analysis, Emotion Recognition, Voice Processing, Tampering Detection, Chatbot, Digital Forensics, Multi-modal System

## I. INTRODUCTION

The increasing prevalence of digital technologies and multimedia communication has significantly transformed the nature of crime and its investigation. Modern crimes often involve diverse forms of evidence such as images, videos, audio recordings, and textual interactions, making the investigation process more complex and time-sensitive. Traditional crime analysis methods largely depend on manual interpretation, which can be time-consuming, error-prone, and insufficient for handling large volumes of heterogeneous data. This creates a growing need for intelligent systems that can assist law enforcement agencies and support victims through automated and real-time analysis.

Recent advancements in Artificial Intelligence (AI), particularly in computer vision, natural language

processing, and audio signal processing, have enabled the development of systems capable of extracting meaningful insights from complex data sources. However, many existing solutions focus on isolated functionalities such as facial recognition, speech processing, or chatbot assistance, without providing a unified framework that integrates these capabilities. Moreover, current systems often lack contextual interpretation, user-centric design, and real-time responsiveness, limiting their effectiveness in practical crime investigation scenarios.

In response to these challenges, this paper presents a Crime AI Support System, a multi-modal backend framework designed to assist in crime analysis and victim support. The proposed system integrates multiple AI-driven modules, including facial emotion recognition, voice-based emotion and speech analysis, media tampering detection, and an intelligent chatbot. These components work collaboratively to analyze multimedia inputs, detect emotional states, verify the authenticity of digital evidence, and provide contextual assistance to users.

The system is implemented using a Flask-based backend architecture, enabling efficient handling of real-time data through RESTful APIs. Unlike conventional systems that merely process raw inputs, the proposed framework performs advanced analysis and generates structured outputs such as emotional insights, risk levels, evidence verification results, and conversational guidance. This reduces the cognitive burden on users and enhances decision-making capabilities for both investigators and victims.

A key feature of the system is its hybrid intelligence approach, combining rule-based logic with AI-driven models to ensure reliability even in resource-

constrained or offline environments. Additionally, the modular design of the backend allows seamless integration with frontend interface

## II. RELATED WORK

The application of Artificial Intelligence in crime analysis and investigation has gained significant attention in recent years, driven by the increasing availability of multimedia data and advancements in machine learning techniques. Early crime support systems primarily relied on manual investigation methods and rule-based digital tools, which were limited in their ability to process large volumes of heterogeneous data such as images, audio recordings, and textual information. While these systems provided basic assistance, they lacked automation, scalability, and real-time analytical capabilities required for modern crime scenarios.

With the emergence of deep learning, several studies have explored the use of computer vision techniques for facial analysis and emotion recognition. Models based on Convolutional Neural Networks (CNNs) have demonstrated high accuracy in detecting facial expressions, enabling applications in surveillance, behavioral analysis, and security systems. Frameworks such as DeepFace have further simplified the integration of pre-trained models into real-time applications. However, most existing implementations focus solely on emotion detection without incorporating contextual interpretation or linking emotional states to crime-related insights, limiting their practical utility in investigative environments.

In the domain of audio analysis, speech recognition systems and voice emotion detection have been widely studied. Techniques involving Mel Frequency Cepstral Coefficients (MFCC), spectral features, and machine learning classifiers have shown effectiveness in identifying emotional cues from speech signals. Additionally, speech-to-text systems using cloud-based APIs have enabled automated transcription of audio evidence. Despite these advancements, existing solutions often treat speech recognition and emotion analysis as separate tasks and rarely integrate them into a unified

framework that extracts both semantic and emotional information simultaneously.

Digital media forensics has also seen significant progress, particularly in the detection of tampered images, videos, and audio files. Methods such as Error Level Analysis (ELA), metadata inspection, noise inconsistency detection, and temporal analysis have been widely used to identify manipulated content. While these techniques are effective individually, many existing systems implement them in isolation, lacking a comprehensive multi-modal approach that combines multiple forensic methods to improve reliability and confidence in tampering detection.

Conversational AI and chatbot systems have been increasingly adopted for user assistance in various domains, including customer support and healthcare. Recent developments in natural language processing have enabled the creation of intelligent chatbots capable of understanding context and generating human-like responses. However, in the context of crime support, most chatbot systems either focus on general-purpose interaction or lack domain-specific knowledge related to legal guidance, emotional support, and emergency response. Furthermore, reliance on cloud-based AI models often limits their availability in offline or resource-constrained environments.

From a system architecture perspective, lightweight backend frameworks such as Flask have been widely used for developing RESTful APIs due to their modularity, scalability, and ease of integration. Existing implementations demonstrate the effectiveness of Flask-based systems for deploying AI models and handling multimedia data

However, many lack a unified architecture that integrates multiple AI modules, structured data processing, and real-time response generation within a single platform.

In contrast to existing approaches, the proposed system introduces a unified multi-modal framework that integrates facial emotion recognition, voice analysis, media tampering detection, and an intelligent chatbot into a single backend architecture. By combining multiple AI techniques with context-

aware interpretation and user-centric design, the system addresses key limitations of prior work and provides a comprehensive solution for crime analysis and victim support.

### III. SYSTEM WORKFLOW

The proposed Crime AI Support System operates through a structured request–response workflow designed to analyze multimedia inputs and provide real-time insights for crime investigation and victim support. Unlike conventional systems that act as passive data handlers, the proposed backend actively processes input data using multiple AI modules to generate meaningful interpretations, risk assessments, and actionable outputs.

#### 1. System Initialization

The system initializes by configuring the backend environment and loading required AI modules, including facial emotion recognition, voice analysis, tampering detection, and chatbot services. Necessary resources such as model dependencies, file storage directories, and API configurations are prepared to ensure smooth operation. This initialization phase ensures system readiness and reliability for handling diverse input types.

#### 2. User Input and Request Handling

Users interact with the system through a frontend interface, where they can provide different types of inputs such as:

- Image frames (via webcam or upload)
- Audio recordings
- Media files (images, videos, audio)
- Text messages for chatbot interaction

These inputs are transmitted to the backend through RESTful API endpoints in structured formats such as JSON or multipart file uploads. The backend validates the inputs to ensure correctness and compatibility.

#### 3. Data Routing and Module Selection

Based on the type of user input, the backend dynamically routes the request to the appropriate AI module:

- Facial images → Emotion Detection Module
- Audio files → Voice Analysis Module
- Media files → Tampering Detection Module
- Text input → Chatbot Module

This modular routing mechanism enables efficient processing and scalability of the system.

#### 4. AI-Based Data Processing

Each module processes the input data using specialized techniques:

- Facial Emotion Detection: Extracts facial features and predicts emotional states using deep learning models.
- Voice Analysis: Performs speech recognition and extracts acoustic features to identify emotional cues and keywords.
- Tampering Detection: Applies forensic techniques such as Error Level Analysis, metadata inspection, and signal processing to verify media authenticity.
- Chatbot Processing: Generates contextual responses using a hybrid rule-based and AI-driven approach.

#### 5. Contextual Interpretation and Risk Assessment

The outputs from individual modules are further processed to generate contextual insights. Emotional states are mapped to risk levels, voice features are interpreted for stress or distress, and tampering scores are evaluated to determine evidence reliability. This layer transforms raw analytical results into meaningful information for decision-making.

#### 6. Data Aggregation and Report Preparation

The system aggregates results from multiple modules and structures them into a unified format. In addition, a report generation component compiles:

- Victim emotional analysis
- Evidence verification results

- Chatbot interaction summaries

This enables the creation of structured case reports for further investigation.

#### 7. Response Generation and Delivery

Finally, the backend returns a structured JSON response containing:

- Analysis results
- Confidence scores
- Risk levels
- Contextual notes

This response is delivered to the frontend, where it can be visualized or used for further interaction, completing the workflow cycle.

#### A. Backend Processing Logic

Unlike traditional systems, the backend performs intelligent preprocessing and interpretation to enhance usability and effectiveness.

##### 1. Multi-Modal Data Analysis

Different data types (image, audio, text) are processed using specialized AI models to extract relevant features and insights.

##### 2. Context-Aware Interpretation

Outputs such as emotions and tampering indicators are mapped to risk levels and investigative notes, improving decision-making.

##### 3. Error Handling and Robustness

The system detects invalid inputs, missing data, or processing failures and returns meaningful error messages to ensure reliability.

#### B. Operational Steps

1. User submits input (image, audio, media, or text)
2. Backend validates and routes the request
3. Relevant AI module processes the input

4. Results are analyzed and contextualized
5. Structured response and reports are generated
6. Output is delivered to the frontend

### IV. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

#### A. System Evaluation

The performance of the proposed Crime AI Support System was evaluated based on accuracy, responsiveness, and robustness across different input types, including facial images, audio recordings, and media files. The system demonstrated consistent performance in detecting facial emotions, analyzing voice characteristics, and identifying potential media tampering. The modular architecture ensured stable operation under multiple concurrent requests, making it suitable for real-time applications in investigative environments.

#### B. Data Interpretability and Insight Generation

Compared to conventional systems that provide isolated outputs, the proposed system enhances interpretability by generating contextual insights such as risk levels, emotional states, and officer guidance notes. The integration of multi-modal analysis enables users to correlate facial expressions, vocal emotions, and textual interactions, thereby improving situational understanding. This approach reduces the need for manual interpretation and supports faster decision-making in critical scenarios.

#### C. Response Efficiency

The system utilizes a lightweight Flask-based backend, resulting in low response latency and efficient processing of multimedia data. Real-time analysis of images and audio is achieved through optimized data handling and selective processing techniques. Additionally, backend-level processing minimizes redundant data transfer, ensuring faster response times and improved system efficiency.

### V. ADVANTAGES AND LIMITATIONS

#### A. Advantages

##### Multi-Modal Intelligence:

The system integrates facial emotion recognition, voice analysis, tampering detection, and chatbot

interaction within a unified framework, providing comprehensive analytical capabilities.

#### Real-Time Decision Support:

Immediate analysis of multimedia inputs enables faster response and improved situational awareness for investigators.

#### Context-Aware Insights:

Outputs are enriched with risk levels and officer notes, transforming raw data into actionable intelligence.

#### Modular and Scalable Architecture:

The RESTful API design allows easy integration with frontend interfaces and future AI modules, ensuring scalability and flexibility.

### B. Limitations

#### 1. Model Dependency:

The accuracy of emotion detection and voice analysis depends on pre-trained models and may vary across different environments and datasets.

#### 2. Environmental Sensitivity:

Facial and voice analysis performance can be affected by factors such as lighting conditions, background noise, and input quality.

#### 3. Limited Forensic Depth:

Tampering detection methods provide indicative results but may not guarantee complete forensic validation in highly sophisticated manipulations.

#### 4. External Library Reliance:

The system depends on third-party libraries and tools, which may affect performance and compatibility over time.

## VI. FUTURE WORK

#### 1. Advanced AI Model Integration:

Incorporation of more robust deep learning models for facial emotion recognition and voice analysis to improve accuracy across diverse environments and datasets.

#### 2. Real-Time Surveillance Integration:

Extension of the system to support live CCTV feeds and continuous monitoring for proactive crime detection and behavior analysis.

#### 3. Enhanced Forensic Analysis:

Implementation of advanced tampering detection techniques such as deepfake detection and frame-level video forensics for improved evidence validation.

#### 4. Intelligent Decision Support System:

Development of predictive models to identify high-risk situations and provide automated alerts to law enforcement authorities.

#### 5. Multilingual and Voice-Based Chatbot:

Expansion of chatbot capabilities to support multiple languages and real-time voice interaction for better accessibility and user engagement.

#### 6. Cloud Deployment and Scalability:

Deployment on cloud platforms to ensure scalability, high availability, and efficient handling of large-scale data and concurrent users.

## VII. CONCLUSION

This paper presented a multi-modal, AI-driven Crime AI Support System designed to assist in modern crime investigation and victim support. By integrating facial emotion recognition, voice analysis, media tampering detection, and an intelligent chatbot within a unified backend framework, the system transforms raw multimedia inputs into actionable insights. The use of a lightweight Flask architecture ensures modularity, scalability, and seamless integration with frontend applications.

Unlike conventional systems that analyze data in isolation, the proposed system emphasizes contextual interpretation and real-time responsiveness by combining multiple AI techniques. The inclusion of risk assessment, emotional analysis, and forensic validation enhances the overall decision-making process for investigators. Additionally, the chatbot component provides interactive assistance, improving user engagement and accessibility.

Overall, the system demonstrates how a unified, backend-centric AI framework can bridge the gap between raw data processing and intelligent crime analysis. The proposed approach lays a strong foundation for future advancements in real-time surveillance, predictive analytics, and intelligent decision support systems, contributing toward more efficient, reliable, and technology-driven crime investigation solutions.

#### REFERENCES

- [1] T. Kopalidis, V. Solachidis, N. Vretos, and P. Daras, "Advances in Facial Expression Recognition: A Survey of Methods, Benchmarks, Models, and Datasets," *Information*, vol. 15, no. 3, p. 135, Feb. 2024, doi: 10.3390/info15030135.
- [2] H. Lian, C. Lu, S. Li, Y. Zhao, C. Tang, and Y. Zong, "A Survey of Deep Learning-Based Multimodal Emotion Recognition: Speech, Text, and Face," *Entropy*, vol. 25, no. 10, p. 1440, Oct. 2023, doi: 10.3390/e25101440.
- [3] Z. Huang et al., "A Study on Computer Vision for Facial Emotion Recognition," *Scientific Reports*, vol. 13, p. 8425, May 2023, doi: 10.1038/s41598-023-35446-4.
- [4] C. Barhoumi and Y. BenAyed, "Real-Time Speech Emotion Recognition Using Deep Learning and Data Augmentation," *Artificial Intelligence Review*, vol. 58, 2024.