Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Enhanced CNN and SVM with Adaptive Modality Switching and Audio-Based Video Summarization for Real-Time Agricultural Intrusion Detection

B Priyanka[1], M Kezia Joseph[2], B Rajendra Naik[3]
1Assistant Professor, Department of Electronics and Communication Engineering, Sreenidhi Institute of Science and Technolog, Email: priyanka.b@sreenidhi.edu.in 2Professor, Department of Electronics and Communication Engineering, Stanley College of Engineering and Technology for Women
3Professor, Department of Electronics and Communication Engineering, University College of Engineering Osmania University

ARTICLE INFO

Received: 15 Dec 2024

Revised: 20 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

Smart intrusion detection in agriculture involves the use of IoT, AI, and sensor-based technologies to monitor fields for unauthorized human and animal activity. Advanced AI models enhance detection accuracy, reducing false alarms and improving response efficiency. The integration of edge computing and cloud-based analytics ensures rapid data processing, making intrusion detection systems more effective and reliable in modern agricultural security. Traditional security systems rely on either video-only or audio-only detection, and struggle in low-light conditions due to the absence of adaptive switching mechanisms. They typically lack event-based video summarization, leading to prolonged, redundant footage storage. Additionally, these systems are high-end and cloud-based, requiring significant computational resources. They are generally designed for broad security applications rather than specialized use cases. The proposed agriculture intrusion detection system consists of an audio based video summarization(ABVS) intrusion detection using an external sensor and day light conditions. The audio based video intrusion system consists of hybrid model with a yolov4-tiny model integrated to a machine learning based audio model which uses MFCCs, Delta and Delta-Delta MFCCs, Chroma features, SMOTE, SpecAugment techniques to improve detection accuracy. An integrated convolution neural network (CNN) model and an integrated SVM model based system termed as “audio-video intrusion detection system with adaptive switching” based on light conditions was implemented with accuracy of 98% and 94%. The comparative t-SNE plot insights were logically used to improve the accuracy of the model as well as the noise augmentation technique with MFCC plus spectral features assisted in achieving the highest efficiency.

Keywords: Precision Agriculture, Sensor Network, Soil Monitoring, Weather Monitoring, Real-Time Data, Agricultural Productivity, Sustainability

INTRODUCTION

Agriculture plays a vital role in global food security, making it essential to safeguard farms from various threats, including human trespassing, wildlife intrusion, and theft. Traditional security measures, such as fences and manual surveillance, are often insufficient, especially in large farmlands where continuous monitoring is impractical. Advancements in technology have led to the integration of smart surveillance systems in agriculture, utilizing sensors, cameras, and artificial intelligence to enhance security. However, these conventional systems still face challenges such as poor visibility in low-light conditions, high false alarm rates due to environmental noise, and the need for continuous human intervention. Intrusion detection in agricultural settings requires a reliable, automated approach that can operate efficiently in diverse environmental conditions. Unlike urban security systems, agricultural surveillance must account for factors such as open-field settings, minimal infrastructure, and unpredictable weather conditions. Wildlife movement, rustling, and unauthorized human access pose significant risks to crop yield and livestock safety. Moreover, traditional video surveillance systems often struggle in nighttime conditions, while audio-based detection alone may be affected by background noise such as wind, machinery, and animal sounds. With the

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

rise of smart farming technologies, integrating intelligent security solutions has become a growing necessity. Modern intrusion detection systems leverage machine learning, Internet of Things (IoT) devices, and advanced data processing techniques to provide real-time threat assessment. These systems aim to enhance accuracy, reduce false alarms, and ensure timely responses to potential security threats. As agricultural lands continue to expand and adopt automation, the need for efficient, cost-effective, and intelligent surveillance solutions becomes increasingly important. Smart security systems tailored for agricultural environments can significantly improve farm protection, minimize losses, and contribute to sustainable farming practices.

Agricultural security is a growing concern due to increasing threats from wildlife, trespassers, and theft, which can result in significant financial and resource losses. Traditional security systems primarily rely on either video-only or audio-only detection methods, which have inherent limitations. Video-based surveillance often struggles in low-light conditions, while audio-based systems may generate false alarms due to environmental noise. Moreover, most conventional systems operate on high-end cloud infrastructure, requiring continuous internet connectivity and high computational power, making them impractical for rural and remote agricultural settings. These limitations highlight the need for a more adaptive, efficient, and cost-effective solution for intrusion detection in farms and agricultural lands. A research study presents an IoT-based farm intrusion detection system designed to monitor agricultural fields for unauthorized intrusions, including human trespassers and wild animals. The system integrates Passive Infrared (PIR) motion sensors, image processing cameras, and acoustic sensors to detect movements and classify them using machine learning algorithms. When an intrusion is detected, real-time alerts are sent to the farmer via SMS or mobile notifications. The study highlights the effectiveness of IoT in improving farm security, reducing losses caused by theft and animal intrusions. It also discusses challenges such as false positives, environmental interferences, and system power consumption. The research concludes that an AI-integrated IoT system can significantly enhance security in remote agricultural fields by enabling real-time monitoring and automated deterrence mechanisms [1].

Hussain, et al. focuses on an IDS framework that enhances agricultural security by implementing advanced feature selection techniques to improve intrusion detection accuracy. The study employs a dataset of various intrusion events, using machine learning models such as Support Vector Machines (SVM), Random Forest, and Neural Networks for classification. It highlights how feature selection techniques like Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) optimize detection by reducing redundant data. The research findings indicate that machine learning-based IDS improves detection rates and reduces false alarms in smart farms. Additionally, the paper discusses the importance of integrating IDS with edge computing to reduce latency in alert responses. The study emphasizes that while IoT devices enhance agricultural automation, security vulnerabilities must be addressed through optimized IDS solutions [2]. A study proposes an Edge-AI-based IDS for real-time detection of human and animal intrusions in smart agricultural environments. The research discusses the limitations of cloud-based systems, including high latency and bandwidth constraints, and advocates for edge computing as a more efficient alternative. The proposed system utilizes low-power AI-driven microcontrollers that process sensor and camera data at the edge, allowing rapid detection of intrusions without reliance on cloud processing. The study demonstrates that edge-based IDS reduces response time and improves energy efficiency, making it a viable solution for large-scale smart farms. Additionally, the research highlights how deep learning models can be deployed on edge devices to classify intrusions with high accuracy. The paper concludes that integrating edge computing with IoT-based security systems can enhance farm surveillance and reduce damage caused by intrusions [3]. A system by Kayes, et al explores cybersecurity threats and intrusion detection in Agriculture 4.0, where IoT networks play a crucial role in monitoring farm activities. The study provides a comparative analysis of various machine learning algorithms for detecting unauthorized access in agricultural fields. It discusses how traditional methods, such as motion-based alarms, fail to distinguish between legitimate and harmful activities, leading to false alerts. The research proposes an IDS model that leverages Convolutional Neural Networks (CNN) for image-based intrusion detection and Long Short-Term Memory (LSTM) networks for analyzing time-series movement data. Experimental results demonstrate that the hybrid approach outperforms conventional security measures, offering greater accuracy in intrusion detection. The study concludes that AI-enhanced IDS solutions can significantly improve security in smart farming ecosystems, making farms less susceptible to external threats [4]. To address these challenges, we propose a Smart Agricultural Intrusion Detection System that leverages Adaptive Audio-Video Fusion for enhanced accuracy and real-time threat detection as shown in figure 1.

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The presented system employs a PIR (Passive Infrared) sensor-based mode switching mechanism, ensuring optimal detection based on environmental conditions. During the day, video surveillance using YOLOv4-Tiny is

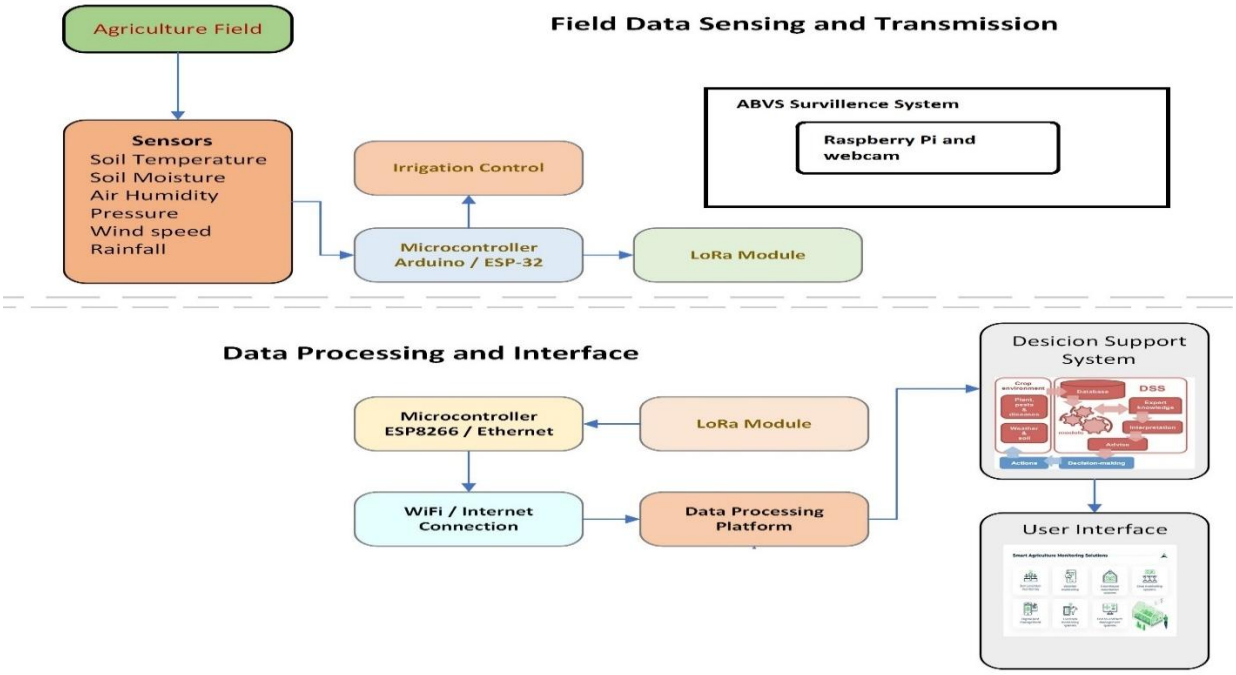

Figure 1. System Overview

prioritized, while at night, the system relies on audio-based classification using a Convolutional Neural Network (CNN) for improved intrusion detection in low-light scenarios. Additionally, the system incorporates Audio-Based Video Summarization (ABVS), which extracts key frames from recorded videos based on detected audio events, reducing redundant footage and enhancing efficiency.

Unlike traditional cloud-based solutions, the proposed system is designed for Raspberry Pi-based Edge Computing, making it cost-effective, power-efficient, and suitable for real-time deployment in agricultural fields. By integrating multiple modalities and leveraging advanced machine learning techniques, this system provides a robust, scalable, and intelligent intrusion detection mechanism tailored specifically for smart agricultural applications. The fusion of audio and video analysis ensures higher accuracy, reduces false alarms, and enhances farm security, making it a reliable and practical solution for modern agricultural surveillance.

RELATED WORK

Integrating audio into video surveillance systems enhances the detection and classification of intrusions in agricultural fields by providing additional sensory data. Audio sensors can capture sounds associated with human or animal activity, such as footsteps, vocalizations, or machinery noise, complementing visual data to improve accuracy. Advanced systems employ machine learning algorithms to analyze both audio and video inputs, enabling real-time identification of potential threats and reducing false positives. This multimodal approach ensures a more robust monitoring system, allowing for timely alerts and responses to unauthorized intrusions, thereby safeguarding crops and property. FarmBeats is an IoT-based smart farming platform that integrates multiple sensors and AI-driven analytics to enhance agricultural productivity and security. A research describes how FarmBeats utilizes low-cost cameras, soil sensors, and drones to monitor fields in real time. While primarily focused on precision agriculture, the platform also includes an intrusion detection module that employs computer vision to identify unauthorized human and animal movements. The research highlights the role of cloud computing in storing and analyzing large volumes of sensor data, enabling predictive analytics for farm security. Additionally, the study discusses challenges such as connectivity issues in remote areas and the need for robust data encryption to prevent cyber threats. The paper

Research Article

concludes that FarmBeats represents a scalable solution for modern farms, combining security, automation, and data analytics to enhance agricultural efficiency [5].

The research work by J.Patel, et al. proposes a system that utilizes wireless sensors to analyze video clips from a gathered dataset, creating an animal intrusion detection mechanism. The system employs deep learning algorithms to process the video data, enabling accurate identification of animal intrusions in agricultural fields. By leveraging wireless sensor networks, the system facilitates real-time monitoring and detection, providing farmers with timely alerts to prevent potential crop damage caused by animals. The integration of deep learning techniques enhances the system's ability to distinguish between different types of intrusions, thereby improving the overall security and management of agricultural fields [6]. AiProff's PashuTham is an AI-driven solution designed to detect and deter wildlife intrusions in Indian farms. The system integrates advanced AI-based image and video analytics to monitor agricultural fields, identifying potential threats from wildlife. By analyzing patterns and behaviors, PashuTham provides real-time alerts to farmers, enabling them to take proactive measures to protect their crops. The solution emphasizes economic viability and ecological sustainability, ensuring that deterrent methods are effective without causing harm to the animals. This approach not only safeguards agricultural produce but also maintains the ecological balance, promoting coexistence between farming activities and wildlife [7].

A model that combines the YOLOv5 object detection algorithm with IoT technology to detect and track stray animals in agricultural fields. The system processes real-time video feeds to identify the presence of animals, achieving an accuracy rate above 96%. Upon detection, the integrated alert system notifies farmers instantly, allowing them to respond promptly to potential threats. The study highlights the effectiveness of this approach in safeguarding crops and ensuring the safety of both farmers and animals. Future enhancements may include expanding the variety of species detectable and optimizing the alert mechanisms for faster notifications [8]. System proposed by an autor utilizes image processing techniques for real-time animal detection in agricultural fields. The setup includes a camera capturing live video streams, which are then analyzed using deep learning algorithms to identify animal intrusions. Upon detection of a potential threat, the system activates a repellent mechanism, such as a buzzer emitting predator sounds, to deter the animal without causing harm. Additionally, the system monitors water levels, transmitting data to nearby storage systems if overloads are detected. This comprehensive approach aims to protect crops from animal damage while ensuring efficient water management [9].

Implementing YOLOv4 (You Only Look Once version 4) for real-time object detection on edge devices like the Raspberry Pi offers a low-cost, highly efficient solution for many applications, such as smart agriculture, surveillance, and robotics. YOLOv4 is known for its superior accuracy and speed compared to other object detection algorithms, making it suitable for deployment on resource-constrained devices. This model uses a single neural network to predict bounding boxes and class probabilities for objects, which enables high accuracy in real-time object detection tasks [10]. However, deploying YOLOv4 on edge devices like Raspberry Pi presents challenges due to limited computational resources, such as processing power and memory. To overcome this, optimizations are necessary, including the use of lightweight versions like YOLOv4-Tiny, which significantly reduces the number of parameters and computational complexity, enabling faster processing without a significant loss in detection accuracy. YOLOv4-Tiny can run on devices with limited resources such as the Raspberry Pi, making it ideal for applications where real-time object detection is required but resources are constrained [11]. For applications such as smart agriculture, where detecting intrusions from humans or animals is critical, YOLOv4-Tiny can be deployed on the Raspberry Pi to provide low-cost solutions for surveillance and monitoring. In this context, it can detect both human and animal movements in agricultural fields, preventing theft, crop damage, and ensuring the security of farm assets. YOLOv4's implementation enables the system to perform object detection efficiently, even with the resource limitations of a Raspberry Pi. In scenarios where real-time monitoring of farm activities is crucial, such as monitoring unauthorized access or animal intrusions, YOLOv4 can be optimized for performance using edge computing solutions like the Raspberry Pi, which reduces the need for cloud-based processing and enables instant responses [12].

OBJECTIVES

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

This proposed study presents a hybrid intrusion detection system for agricultural security, integrating YOLOv4-Tiny for video analysis during the day and audio classification for night-time detection, guided by a PIR sensor. The system leverages Audio-Based Video Summarization (ABVS) to categorize video frames based on audio features, ensuring efficient event monitoring. Designed for edge deployment on Raspberry Pi, it offers a low-power, real-time solution optimized for noisy farm environments using MFCC and spectral features for robust audio classification.

Key Contributions:

•Hybrid Detection: Combines YOLOv4-Tiny for video-based intrusion detection during the day and audio classification at night using a PIR sensor.

•ABVS Integration: Implements Audio-Based Video Summarization (ABVS) to enhance video analysis by categorizing frames based on audio cues.

•Efficient Edge Deployment: Designed for Raspberry Pi, ensuring low-power, real-time processing suitable for remote agricultural applications.

•Noise-Augmented Audio Classification: Utilizes MFCC and spectral features for improved audio-based intrusion detection in noisy farm environments

SYSTEM DESIGN AND IMPLEMENTATION

System Architecture: Conventional security systems typically rely on either video-only or audio-only detection, whereas an Adaptive Audio-Video Fusion approach intelligently integrates both modalities to enhance accuracy. Regarding light sensitivity, conventional systems often struggle in low-light conditions, reducing their effectiveness at night. In contrast, an adaptive system utilizes a PIR sensor-based mode-switching mechanism, prioritizing video analysis during the day and audio-based detection at night for optimal intrusion identification. Most traditional systems lack event-based video summarization, resulting in prolonged and often redundant footage. However, an adaptive system incorporates Audio-Based Video Summarization (ABVS) to extract key frames based on relevant audio cues, ensuring concise and meaningful summaries. In terms of deployment, conventional security solutions are often high-end and cloud-based, requiring substantial computational resources. Conversely, an adaptive system is optimized for Raspberry Pi (Edge Computing), enabling efficient real-time intrusion detection with lower hardware requirements. While traditional security systems cater to general security applications, an adaptive approach is specifically designed for Agricultural Intrusion Detection, addressing unique challenges such as wildlife threats, farm security, and environmental conditions.

Components of IoT based Intrusion Detection System: IoT-based Intrusion Detection Systems (IDS) play a crucial role in enhancing security in agricultural fields by continuously monitoring for unauthorized human and animal intrusions. These systems leverage a combination of sensors, cameras, communication modules, and control units to detect, analyze, and respond to potential threats in real time. Motion and acoustic sensors provide primary detection, while cameras with AI-driven image processing enhance accuracy. Communication modules such as WiFi, Zigbee, LoRa, and GSM enable remote monitoring, and control units process data to trigger automated deterrent responses. This integration ensures efficient, cost-effective, and scalable security solutions for modern smart agriculture.

Sensors:

Sensors form the backbone of IoT-based Intrusion Detection Systems (IDS) in agriculture, enabling real-time monitoring of fields to detect unauthorized activities. Various types of sensors are deployed, each serving a specific purpose:

- Motion Sensors: Passive Infrared (PIR) sensors detect infrared radiation emitted by warm objects, such as humans and animals, within a specified range. When movement is detected, they trigger alerts, allowing farm owners to take preventive measures. PIR sensors are widely used due to their low power consumption and cost-effectiveness.
- Acoustic Sensors: By capturing sound patterns, acoustic sensors distinguish between different sources of noise, such as human voices, barking dogs, or animal calls. Advanced systems utilize machine learning algorithms to analyze sound waves and classify them accurately. For example, a sudden increase in noise levels from a specific direction may indicate an intrusion, prompting further investigation.

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Cameras:

- Cameras play a crucial role in IoT-based IDS by providing visual confirmation of intrusions. Unlike traditional surveillance systems, modern agricultural security cameras are equipped with image processing and artificial intelligence (AI) capabilities to enhance detection accuracy.
- AI and Machine Learning Integration: Advanced systems integrate deep learning algorithms to analyze video footage and identify specific intruders. These models can distinguish between human and animal movements, reducing false alerts caused by irrelevant objects like tree branches moving in the wind. For example, an AI-powered system trained on animal behavior patterns can recognize a wild boar or a deer approaching the crops and send an immediate alert to the farmer.
- Live Streaming and Cloud Storage: IoT-enabled cameras can live stream footage to farmers' mobile devices, enabling real-time monitoring. They can also store video clips on cloud-based systems for future analysis. In case of an intrusion, these recordings provide valuable evidence for security or legal actions.

Communication Modules:

Communication modules are essential for transmitting data collected by sensors and cameras to a centralized processing unit or cloud-based system. These modules enable real-time alerts and allow farmers to monitor their fields remotely. The choice of communication technology depends on factors such as range, power consumption, and data transmission speed. In this work WiFi is used for communication as audio and images need to be transmitted and used for sending alerts to the User.

Control Units:

The control unit acts as the brain of an IoT-based Intrusion Detection System, processing data from multiple sensors and cameras to detect and classify intrusions accurately. It ensures the efficient execution of intrusion detection algorithms, manages alert notifications, and triggers deterrent mechanisms.

- Processing Data from Sensors and Cameras: The control unit collects raw data from motion, vibration, and acoustic sensors, as well as image and video feeds from cameras. It processes this data using edge computing or cloud-based systems, depending on the deployment model. Edge computing is preferred for real-time applications as it processes data locally, reducing latency and improving response times.
- Intrusion Detection Algorithms: Advanced control units utilize machine learning algorithms and artificial intelligence (AI) to differentiate between normal activities and genuine threats. For example, convolutional neural networks (CNNs) are used for image recognition, while support vector machines (SVMs) and decision trees help classify sensor data.
- Automated Responses and Alerts: When an intrusion is detected, the control unit triggers automated deterrent mechanisms, such as activating alarms, turning on floodlights, or deploying acoustic deterrents to scare off animals. Simultaneously, alerts are sent to farmers via SMS, mobile apps, or email so they can take immediate action. The control unit ensures the efficient operation of an IoT-based IDS, enhancing security, reducing false alarms, and providing real-time protection against potential threats. By integrating these components—sensors, cameras, communication modules, and control units—IoT-based Intrusion Detection Systems provide comprehensive, automated, and intelligent security solutions for agricultural fields. This ensures better farm management, reduced losses due to intrusions, and enhanced productivity, making IoT a crucial technology in modern precision agriculture.

METHODOLOGY

The proposed system integrates audio-based and video-based intrusion detection for smart agricultural security. The methodology involves training a CNN model for audio classification and utilizing YOLOv4-Tiny for object detection to filter relevant video frames using Audio-Based Video Summarization (ABVS). Figure 2 shows the system flow chart. The following sections detail the training strategy and methodology used to develop and implement this system. The dataset comprises audio recordings of various agricultural intrusion events, including human intrusions, wildlife presence, and environmental sounds. To ensure robust classification, multiple datasets were integrated. The ESC-50 dataset, a labeled collection of 2,000 environmental audio recordings, each 5 seconds long, was utilized, with relevant animal sound classes selectively chosen for training. Additionally, the Indian Human Accent Dataset from

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Kaggle was incorporated to enhance the system's ability to detect human intrusions, ensuring accurate recognition of region-specific speech patterns. This diverse dataset enables effective benchmarking and classification of environmental and intrusion-related sounds in agricultural settings.

Audio classification in an agricultural intrusion detection system requires robust preprocessing techniques to ensure high accuracy and adaptability to real-world conditions. Various factors such as environmental noise, variations in sound characteristics, and class imbalances must be addressed to improve model performance. The preprocessing pipeline includes steps such as audio resampling, feature extraction using MFCCs and additional spectral features, data balancing using SMOTE, and data augmentation techniques to enhance the robustness of the classification model. These steps ensure that the deep learning model effectively differentiates between intrusion sounds, such as human speech and animal noises, and non-intrusive environmental sounds.

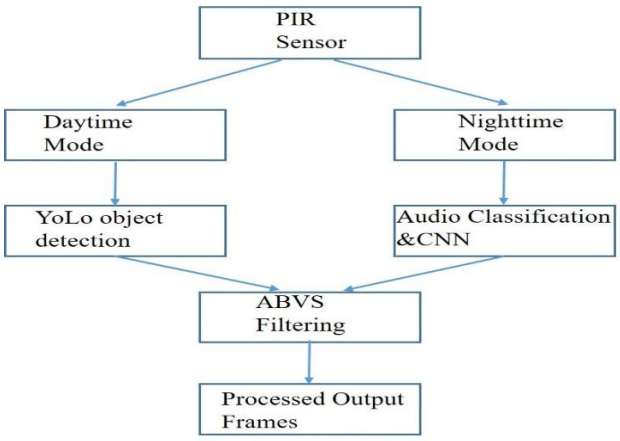

Fig 2. System flow chart

To begin with, all audio recordings are resampled to a standard sampling rate of 16,000 Hz. This step is crucial as different recordings may have been captured at varying sample rates, leading to inconsistencies in feature extraction. Resampling ensures that all input data maintains a uniform temporal resolution, allowing the model to learn patterns consistently across different audio samples. It also helps optimize computational efficiency while preserving essential sound features.

One of the most important features for audio classification is the Mel-Frequency Cepstral Coefficients (MFCCs). MFCCs are widely used in speech and environmental sound classification because they mimic the human auditory perception of sound frequencies. The preprocessing pipeline extracts 40 MFCC coefficients per frame, capturing detailed frequency characteristics that help differentiate between different types of intrusion and non-intrusion sounds. Unlike raw spectrograms, MFCCs emphasize perceptually relevant frequency components, making them ideal for recognizing variations in farm noises, human speech, and wildlife presence.

To further improve classification accuracy, additional spectral features are extracted from the audio signals. These include Delta and Delta-Delta MFCCs, which measure the rate of change in MFCC values over time. These features help the model understand the dynamic properties of sounds, making it more effective at recognizing sequences such as human speech patterns or the distinct movement-based noises made by animals. Spectral Contrast is another important feature that captures the difference between peaks and valleys in the frequency spectrum. This helps distinguish between sounds that may have similar energy distributions but belong to different classes, such as distinguishing between rustling leaves and actual animal movement. Additionally, Chroma features are extracted to represent tonal information based on 12 pitch classes, which are particularly useful for differentiating human speech from background environmental sounds.Figures 3-5 show the various plots of MFCC Extracted from various classes. Figure 6 shows the MFCC Spectrogram of all the classes.

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

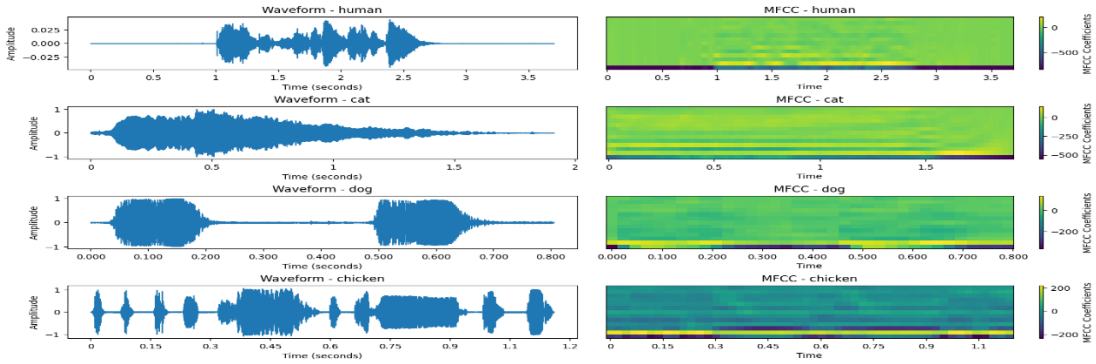

Fig 3. MFCC Extracted plot of human, cat,dog and chicken

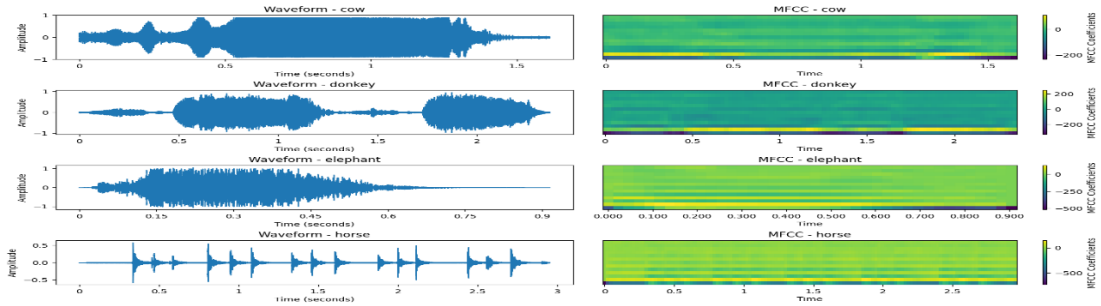

Fig 4. MFCC Extracted plot of cow, donkey, elephant and horse

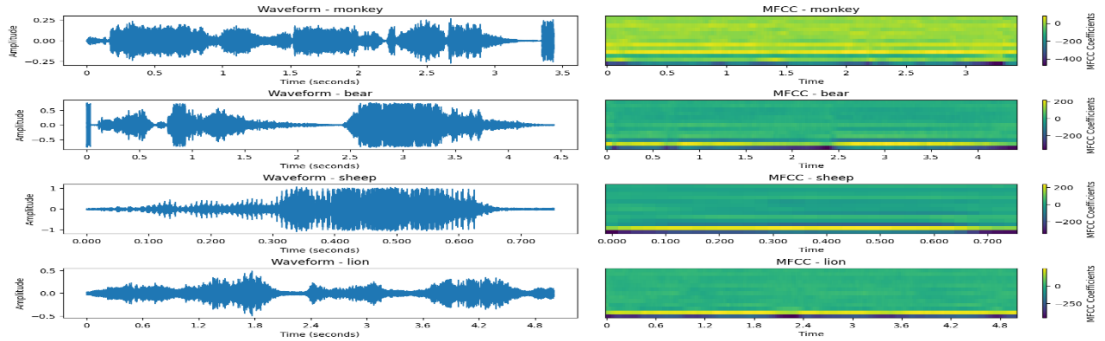

Fig 5. MFCC Extracted plot of monkey,bear,sheep and lion

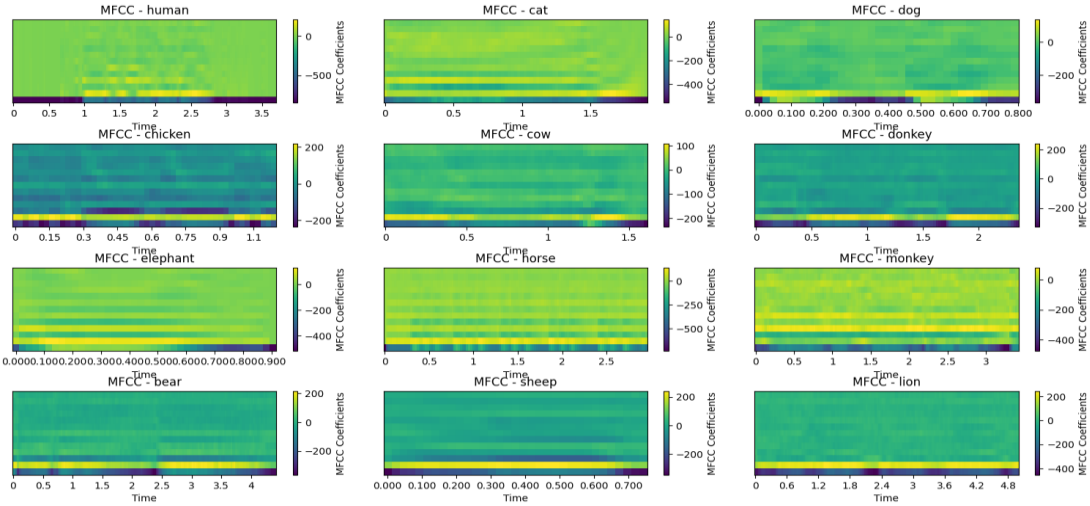

Fig 6. MFCC Spectrogram

One of the challenges in real-world intrusion detection is class imbalance, where certain intrusion events, such as

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

specific animal noises or human voices, occur far less frequently than general environmental sounds. To address this issue, Synthetic Minority Over-sampling Technique (SMOTE) is applied. SMOTE generates synthetic training samples for underrepresented classes by interpolating between existing minority class samples. Unlike simple oversampling techniques that duplicate minority samples, SMOTE creates new data points that maintain the original class distribution while introducing variability, preventing overfitting. This ensures that the model does not become biased toward the majority class, improving its ability to detect rare but critical intrusion events.

In addition to balancing the dataset, data augmentation techniques are applied to improve the model's robustness in real-world environments. Agricultural settings often have unpredictable noise conditions, such as wind, rain, and machinery sounds, which can interfere with accurate classification. To simulate these conditions and enhance generalization, SpecAugment techniques such as random time masking and frequency masking are used. Time masking randomly removes small segments of the audio signal, forcing the model to learn features that are not dependent on specific time intervals. This helps the model become resilient to interruptions or missing parts in an audio recording. Frequency masking, on the other hand, removes certain frequency bands in the spectrogram representation, ensuring that the model does not rely solely on specific frequency ranges and can adapt to variations in sound pitch. These augmentations help the model learn more robust and invariant features, ultimately leading to improved accuracy in noisy environments.

By implementing this comprehensive preprocessing pipeline, the system ensures that audio data is well-prepared for classification. The use of MFCCs and spectral features provides a rich representation of the sound signals, while SMOTE effectively addresses class imbalance, ensuring that all intrusion categories are equally represented. The inclusion of SpecAugment further enhances the model's resilience to environmental noise, making it highly effective for agricultural intrusion detection in real-world conditions.

Model Architecture & Training for Audio-Based Intrusion Detection

The Convolutional Neural Network (CNN) developed for agricultural intrusion detection is designed to process Mel-Frequency Cepstral Coefficients (MFCCs) extracted from audio signals. The architecture consists of convolutional layers, which capture spatial patterns within the MFCCs, identifying key frequency components that distinguish between different intrusion sounds, such as human presence, wildlife activity, and environmental disturbances. Multiple convolutional layers are stacked to extract hierarchical features, enhancing the model's ability to recognize complex audio patterns.

To improve training stability and prevent overfitting, batch normalization and dropout layers are incorporated. Batch normalization ensures that activations remain within a stable range, leading to faster convergence, while dropout layers randomly deactivate neurons during training, making the model more resilient to noise and variations in audio data. The final extracted features are passed through fully connected layers, which classify the audio input by mapping the high-level feature representations to different intrusion categories.

The training process is optimized using the Adam optimizer, which dynamically adjusts learning rates for efficient weight updates. Since the problem involves multi-class classification, the categorical crossentropy loss function is used to ensure proper learning of class probabilities. To ensure that the model generalizes well to different audio samples, 5-fold cross-validation (K=5) is applied. This method splits the dataset into five subsets, training on four while validating on the fifth, cycling through each subset for validation in different iterations. This approach helps assess model robustness and prevents overfitting to specific samples.

Training is conducted over 50 epochs, with an early stopping mechanism that halts training if validation loss ceases to improve, preventing unnecessary computations and overfitting. The final trained CNN model is stored in .keras format for deployment in real-time agricultural intrusion detection systems. Performance evaluation is conducted using accuracy, precision, recall, and F1-score, ensuring the model effectively differentiates between intrusion and non-intrusion events, making it suitable for real-world applications.

Object Detection Model: YOLOv4-Tiny for Intrusion Detection YOLOv4-Tiny is selected for object

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

detection due to its lightweight architecture, making it highly suitable for real-time deployment on edge devices like the Raspberry Pi. Unlike traditional object detection models, YOLOv4-Tiny offers a balance between detection speed and accuracy, ensuring efficient intrusion identification in agricultural settings. The model is pre-trained on the COCO dataset and fine-tuned with agricultural intrusion-specific images to enhance its ability to detect relevant objects such as humans, animals, and vehicles. Image preprocessing involves resizing frames to 416×416 pixels to match the model's input dimensions and normalizing pixel values to optimize detection performance.

For inference, YOLOv4-Tiny operates on the Darknet framework, ensuring efficient real-time object detection. A confidence threshold of 50% is applied to filter out weak detections, ensuring that only reliable object predictions are considered. Additionally, Non-Maximum Suppression (NMS) is implemented to eliminate redundant bounding boxes, ensuring that a single, precise detection is retained per object. The model specifically tracks intrusion-related objects to distinguish between threats and non-threatening elements in the environment.

Adaptive Audio-Video Fusion & ABVS Implementation The proposed system integrates Adaptive Audio-Video Fusion for intelligent intrusion detection, dynamically switching between audio and video modalities based on environmental conditions. A PIR sensor determines the detection mode by assessing ambient light levels. During daytime, when visibility is high, the system prioritizes video-based detection using YOLOv4-Tiny, ensuring accurate object recognition. At night, when visibility is poor, the system relies on parallel audio classification using both CNN and SVM models, ensuring uninterrupted detection through acoustic signals.

The intrusion detection process begins with audio extraction from video files, followed by the computation of MFCC-based features. These features serve as inputs for both CNN and SVM models, which independently classify the detected sound as either an intrusion or non-intrusion event. Simultaneously, YOLOv4-Tiny analyzes video frames to detect objects associated with potential intrusions. To optimize video storage and processing, Audio-Based Video Summarization (ABVS) is applied, retaining only the most relevant frames. If both CNN and SVM models classify an audio event as an intrusion AND YOLOv4-Tiny detects a relevant object, the corresponding video frame is retained. Frames without significant objects or intrusion-related audio cues are discarded, ensuring that only critical footage is stored for further analysis.

This frame filtering process significantly reduces redundant video storage while maintaining essential security footage. The percentage of retained frames versus total frames is computed to measure ABVS efficiency, optimizing the balance between data retention and computational efficiency. This integrated approach ensures that intrusion detection remains robust, accurate, and resource-efficient, making it highly effective for real-world agricultural security applications.

RESULTS

The system's effectiveness is evaluated using multiple metrics. Audio classification performance is assessed using CNN and SVM models, with accuracy, F1-score, and confusion matrices analyzed to understand misclassifications. YOLOv4-Tiny is evaluated using Precision, Recall, and mean Average Precision (mAP) to measure object detection accuracy. The comparative performance of CNN and SVM helps determine the most effective model for real-time intrusion detection.

Additionally, ABVS effectiveness is quantified by computing the percentage of retained frames versus total frames. The system also measures the accuracy improvement from ABVS, ensuring reduced redundancy while maintaining relevant intrusion evidence. By evaluating both CNN and SVM models in parallel, the system ensures a comprehensive and adaptable approach to intrusion detection, enhancing overall security and efficiency.

The t-Distributed Stochastic Neighbor Embedding (t-SNE) visualization provides a clear representation of how well the CNN model separates different audio classes based on learned features. This technique reduces high-dimensional feature space into a 2D plot, helping us understand the clustering behavior of different categories.

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

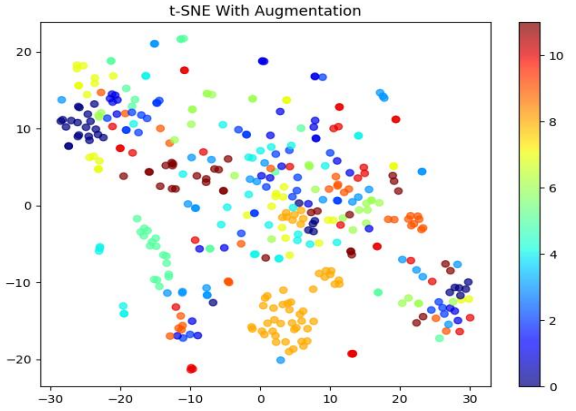

Fig 7. t-SNE plot with noise augmentation

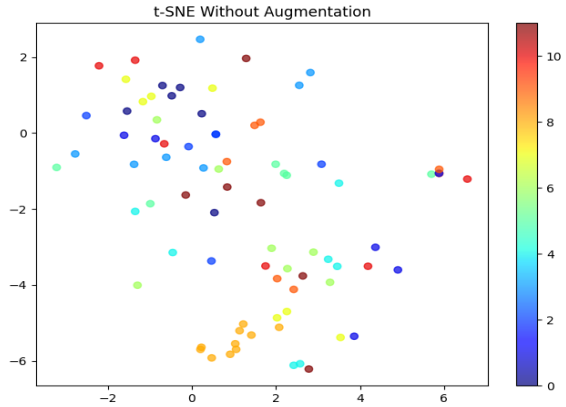

Fig 8. t-SNE plot without noise augmentation

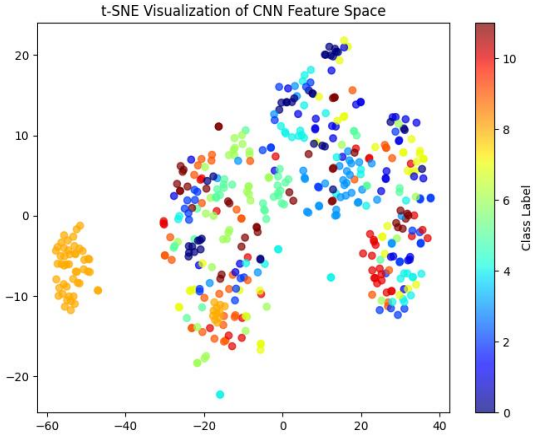

Fig 9. t-SNE visualization of CNN feature space

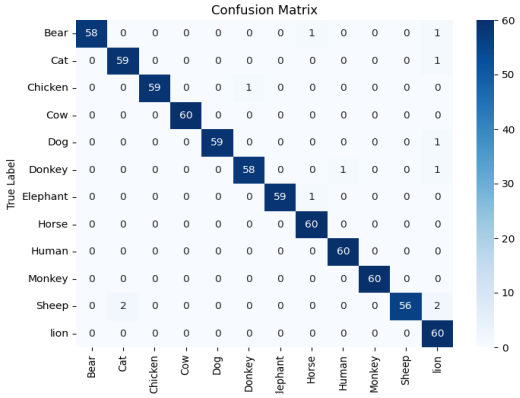

Fig 10. Confusion matrix of CNN feature space

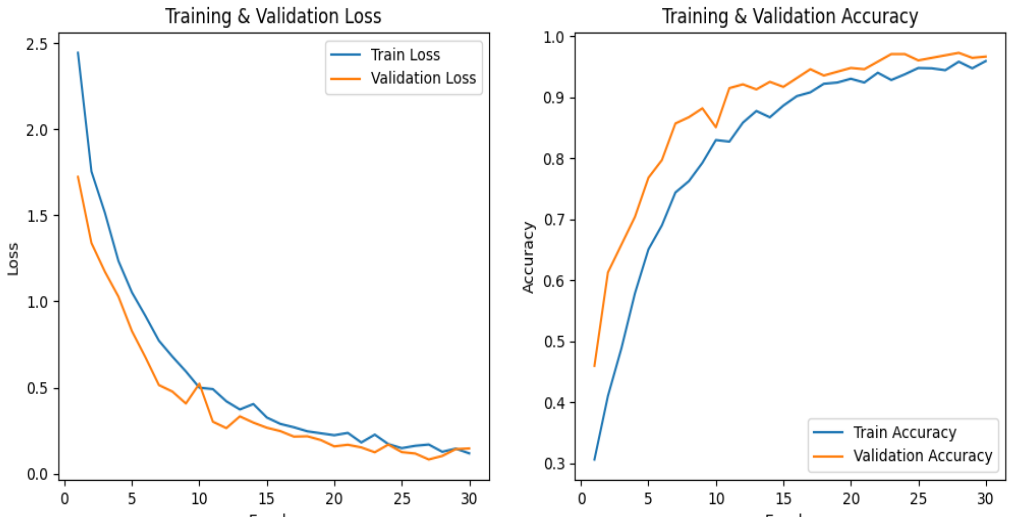

Fig 13 . Training and validation accuracy of CNN model

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

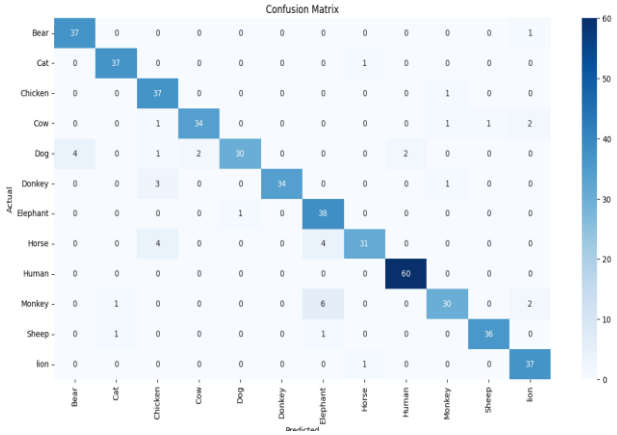

Fig 11. Confusion matrix of SVM

The t-SNE plot reveals that some classes are well-separated, forming distinct clusters. For example, a specific class (represented by orange points in figure 9) is isolated on the left, indicating that the CNN model has successfully learned features that differentiate this class from others. This suggests that the audio patterns in this category are unique, making it easier for the model to classify them correctly. On the other hand, while most classes show a tendency to form clusters, some categories appear closer to each other, hinting at possible confusion. This occurs when different classes share similar spectral characteristics, making them harder to differentiate purely based on their MFCC-based representations. The visualization also highlights the confidence level of the CNN's classifications. Densely packed clusters indicate that the model is highly confident in classifying those instances correctly. These classes likely contain clear, distinguishable patterns that the model has effectively learned.

DISCUSSION

Table 1. Classification Report of CNN

Class	Precision	Recall	F1-Score	Support
Bear	1	0.97	0.98	60
Cat	0.97	0.98	0.98	60
Chicken	1	0.98	0.99	60
Cow	1	1	1	60
Dog	1	0.98	0.99	60
Donkey	0.98	0.97	0.97	60
Elephant	1	0.98	0.99	60
Horse	0.97	1	0.98	60
Human	0.98	1	0.99	60
Monkey	1	1	1	60
Sheep	1	0.93	0.97	60
Lion	0.91	1	0.95	60
Overall Accuracy	**0.98**	—	—	**720**

However, lion sounds have the lowest precision (0.91), indicating that the model occasionally misclassifies other animal sounds as lions, likely due to similarities in growls or roars. Recall, on the other hand, evaluates how well the model identifies actual occurrences of a class. A high recall score suggests fewer false negatives, ensuring that most instances are correctly classified. The sheep class has the lowest recall (0.93), meaning that some sheep sounds were

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

misclassified, possibly as goats or donkeys due to their acoustic resemblance. The F1-score, which balances precision and recall, is consistently high across all classes, confirming the model's reliability. Some classes, such as cow and monkey, achieved perfect precision and recall (1.00), indicating that their sounds have highly distinctive characteristics, making misclassification unlikely. In contrast, classes like lion and sheep exhibit minor misclassification tendencies, suggesting room for further refinement in feature extraction or dataset diversity.

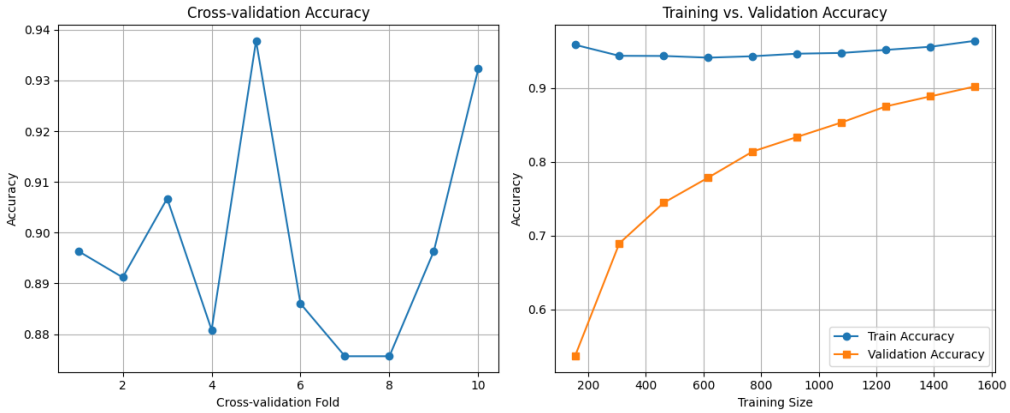

Fig 12. Training and validation accuracy of SVM model

Table 2. SVM Performance metrics

Class	Precision	Recall	F1-Score	Support
Bear	0.95	0.97	0.96	60
Cat	0.98	0.88	0.93	60
Chicken	0.95	0.93	0.94	60
Cow	0.98	0.93	0.96	60
Dog	0.93	0.93	0.93	60
Donkey	0.95	0.93	0.94	60
Elephant	0.97	0.95	0.96	60
Horse	0.88	0.97	0.92	60
Human	0.98	1	0.99	60
Monkey	0.88	0.97	0.92	60
Sheep	1	0.93	0.97	60
Lion	0.94	0.97	0.95	60
Overall Accuracy	**0.947**	—	—	**720**

The SVM results shown in Table 2 and figure 11-12 shows that the model achieved an accuracy of 94.7% (0.9472) after balancing the dataset, demonstrating a significant improvement due to the class balancing strategy, which ensured fair training across all categories. The performance across classes varied, with the Human class achieving the highest accuracy (Precision: 0.98, Recall: 1.00, F1-score: 0.99), meaning the model correctly identified almost all human-related sounds. Similarly, the Sheep class performed well (Precision: 1.00, Recall: 0.93, F1-score: 0.97), though it had a few false negatives. However, some areas still require improvement. The Cat class had high precision (0.98) but lower recall (0.88), indicating misclassification with other categories. The Horse and Monkey classes exhibited lower precision (0.88), suggesting occasional false positives where the model incorrectly predicted these classes.

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 3. Comparison of proposed work and existing work

Ref	Title	Methodology	Dataset	Accuracy	Journal	Year
[13]	Ensemble of Convolutional Neural Networks to Improve Animal Audio Classification	Ensemble of fine-tuned CNNs combined with handcrafted texture descriptors	Bird, bat, and whale audio datasets	Not specified	EURASIP Journal on Audio, Speech, and Music Processing	2020
[14]	Comparison of Pre-Trained CNNs for Audio Classification Using Transfer Learning	3 image and 2 sound'-trained CNNs, namely, and transfer learning is used	UrbanSound8K, ESC-10, and Air Compressor	96.4% for UrbanSound8K, 91.25% for ESC-10, 100% for the Air Compressor	Journal of Sensor and Actuator Networks	2021
[15]	DualDiscWaveGAN-Based Data Augmentation Scheme for Animal Sound Classification	Data augmentation using DualDiscWaveGAN for conditional animal sound generation	North American Bird Species and SK frogs	98.4 on NA Bird and 84.1 on SK frogs	Sensors	2023
[16]	An Accurate and Fast Animal Species Detection System for Embedded Devices	YOLOv2, modified YOLOv2, YOLOv3, and YOLOv4 with deformable convolutional layers for animal detection on Rpi 4.	BCMTI of six Canadian animal species having 138,482 unlabeled images	Outperforms YOLOv2 by 5% accuracy, 12% speed	IEEE Access	2023
[17]	Agrivigilance: A Security System for Intrusion Detection in Agriculture Using Raspberry Pi and OpenCV	RPi with OpenCV for motion detection and object identification using Single Shot detectors and Mobilenets technique of Deep Learning	Not specified	92% accuracy, 100% consistency	International Journal of Scientific & Technology Research	2019
[18]	Application of IOT and machine learning in crop protection against animal intrusion	IoT integration with R-CNN and SSD for animal detection	5 animals classes with 60 images in each class	Proposed R-CNN accuracy 85.22 % Proposed-SSD - 89.32%	Global Transitions Proceedings	2021

Journal of Information Systems Engineering and Management
2025, 10(33s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[19]	A novel Approach for Audio-based Video Analysis via MFCC Features	ABVS employs 40 MFCC audio features, various ML and ML ensemble methods are trained	Urbansound8k dataset	accuracy of 90% and loss of 0.10 on ANN and SVM	Procedia Computer Science	2024
[20]	Animal Sound Classification Using A Convolutional Neural Network	Nesterov-accelerated Adaptive Moment Estimation	875 audio samples for 10 animal classes	Accuracy 75%	2018 3rd International Conference on CSE (UBMK)	2018
[21]	Animal Sound Classification Using Dissimilarity Spaces	F_NN ensemble based on Siamese networks	ESC-50	F_NN + eCNN 88.47 96.03	Applied Sciences	2020
	Proposed work	Modified CNN + SVM with PIR-Based Adaptive Switching & ABVS	ESC-50 (Balanced) and online animal data sets, kaggle human data set	98.0% (CNN), 94.7% (SVM)		2025

Additionally, while the Lion class performed well (Precision: 0.94, Recall: 0.97), some lion sounds were still misclassified. The balanced dataset improved recall across all categories, ensuring better generalization. However, some classes still show slight misclassification tendencies, highlighting the need for further feature engineering and hyperparameter tuning to refine classification boundaries. Table 3 shows the comparioion between various existing works with the proposed model and it is seen that proposed model achieves higher accuracy compared to the previous works.

CONCLUSION AND FUTURE SCOPE

The evaluation of the intrusion detection system demonstrates the effectiveness of CNN and SVM models for audio classification, along with YOLOv4-Tiny for object detection. The CNN model achieved an impressive accuracy of 98%, successfully distinguishing between animal and human sounds with minimal misclassification. The SVM model, after dataset balancing, reached an accuracy of 94.7%, showing significant improvement in classification performance across all classes. While both models performed well, CNN demonstrated higher accuracy and better feature learning, making it the more reliable choice for real-time intrusion detection. The t-SNE visualization further validated CNN's feature extraction capability, revealing distinct class clusters, though some overlap existed for acoustically similar sounds. Misclassifications, particularly in the lion and sheep classes, suggest the need for further refinement in feature extraction and dataset diversity to enhance separation. The YOLOv4-Tiny model's performance, evaluated using Precision, Recall, and mean Average Precision (mAP), ensures robust visual detection in different lighting conditions, complementing the audio-based classification.

Additionally, ABVS improved system efficiency by reducing redundant frames while preserving key intrusion evidence, as indicated by the retained-frame percentage analysis and visualization. The integration of audio-based classification, object detection, and ABVS ensures a comprehensive, adaptive, and efficient intrusion detection system. Further enhancements through hyperparameter tuning, feature engineering, and dataset augmentation can refine classification accuracy and improve system robustness in real-world deployments.

The findings from this study highlight key areas for further enhancement in audio-based intrusion detection. One of

the primary challenges observed was the misclassification of Chicken and Elephant sounds, where precision was lower due to overlap with other classes. Future work can focus on improving precision by refining feature extraction techniques to better capture the distinguishing characteristics of these sounds. Additionally, classes like Dog, Monkey, and Horse exhibited lower recall, suggesting that the model could benefit from a more diverse dataset with varied environmental conditions to improve generalization.

Feature engineering can play a crucial role in enhancing classification performance. Incorporating Spectral Contrast, Chroma Features, and delta MFCCs may help in better differentiating acoustically similar sounds. Moreover, hyperparameter tuning—such as adjusting CNN layers, dropout rates, and learning rates—can refine classification boundaries and reduce misclassifications. Expanding the dataset through augmentation techniques like time-stretching, pitch shifting, or noise injection can further strengthen the model's robustness against real-world variations.

REFRENCES

[1] M. R. Ali, N. S. Sayem, S. Chowdhury, and M. Kamruzzaman, "Farm Intrusion Detection System using IoT," IEEE Xplore, 2022.

[2] M. S. Hossain, M. A. Rahman, and G. Muhammad, "A Novel Intrusion Detection System (IDS) Framework for Internet of Things (IoT) using Feature Selection Method," Journal of Theoretical and Applied Information Technology, vol. 101, no. 21, pp. 5758-5768, 2020.

[3] S. S. Khan, M. A. Jan, M. Alam, and M. Usman, "An Intrusion Detection System for Edge-Envisioned Smart Agriculture," IEEE Access, vol. 9, pp. 167056-167065, 2021.

[4] S. M. Kayes, M. R. Islam, and A. I. Khan, "Cyber Security Intrusion Detection for Agriculture 4.0: Machine Learning Based Solutions," IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 9, pp. 1489-1499, 2021.

[5] D. Vasisht, Z. Kapetanovic, J. Won, X. Jin, and R. Chandra, "FarmBeats: An IoT Platform for Data-Driven Agriculture," in Proceedings of the 14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17), Boston, MA, USA, 2017, pp. 515-529

[6] L. Tan and N. Wang, "Future Internet: The Internet of Things," in 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), Chengdu, China, 2010, pp. V5-376-V5-380.

[7] Y. Liu, X. Y. Ma, L. Shu, G. P. Hancke, and A. M. Abu-Mahfouz, "From Industry 4.0 to Agriculture 4.0: Current Status, Enabling Technologies, and Research Challenges," IEEE Transactions on Industrial Informatics, vol. 17, no. 6, pp. 4322-4334, 2021.

[8] M. M. Rathore, A. Paul, A. Ahmad, B. W. Chen, and W. Ji, "Real-time Intrusion Detection for IoT-based Smart Agriculture," IEEE Internet of Things Journal, vol. 8, no. 4, pp. 2345-2356, 2021.

[9] J. Patel, P. Shah, and R. K. Gupta, "Deep Learning-based Animal Intrusion Detection in Smart Farms," Future Internet Journal, vol. 15, no. 12, pp. 1-15, 2023.

[10] K. Bose, A. Sharma, and N. Verma, "A Hybrid IoT and AI Approach for Real-Time Farm Security," IEEE Internet of Things Magazine, vol. 4, no. 2, pp. 30-37, 2023

[11] J. Patel, P. Shah, and R. K. Gupta, "Animal Intrusion Detection using Deep Learning for Agricultural Fields," IEEE Access, vol. 9, pp. 1250-1255, 2021.

[12] Wildlife Intrusion Detection and Prevention in Farmlands using AI

[13] L. Nanni, Y. M. G. Costa, R. L. Aguiar, R. B. Mangolin, S. Brahnam, and C. N. Silla, "Ensemble of convolutional neural networks to improve animal audio classification," EURASIP Journal on Audio Speech and Music Processing, vol. 2020, no. 1, May 2020, doi: 10.1186/s13636-020-00175-3.

[14] E. Tsalera, A. Papadakis, and M. Samarakou, "Comparison of Pre-Trained CNNs for audio Classification using transfer Learning," Journal of Sensor and Actuator Networks, vol. 10, no. 4, p. 72, Dec. 2021, doi: 10.3390/jsan10040072.

[15] E. Kim, J. Moon, J. Shim, and E. Hwang, "DualDiscWaveGAN-Based Data Augmentation Scheme for animal sound classification," Sensors, vol. 23, no. 4, p. 2024, Feb. 2023, doi: 10.3390/s23042024.

[16] M. Ibraheam, K. F. Li, and F. Gebali, "An accurate and fast animal species detection system for embedded devices," IEEE Access, vol. 11, pp. 23462–23473, Jan. 2023, doi: 10.1109/access.2023.3252499.

[17] S. Angadi and R. Katagall, “AgriVigilance: a security system for intrusion detection in agriculture using Raspberry Pi and OpenCV,” International Journal of Scientific and Technology Research, vol. 8, no. 11, pp. 1260–1267, Nov. 2019, [Online]. Available: https://www.ijstr.org/paper-references.php?ref=IJSTR-1119-24621

[18] K. Balakrishna, F. Mohammed, C. R. Ullas, C. M. Hema, and S. K. Sonakshi, “Application of IOT and machine learning in crop protection against animal intrusion,” Global Transitions Proceedings, vol. 2, no. 2, pp. 169–174, Aug. 2021, doi: 10.1016/j.gltp.2021.08.061.

[19] Sabha and A. Selwal, “A novel Approach for Audio-based Video Analysis via MFCC Features,” Procedia Computer Science, vol. 235, pp. 1512–1521, Jan. 2024, doi: 10.1016/j.procs.2024.04.142.

[20] E. Sasmaz and F. B. Tek, “Animal sound classification using a convolutional neural network,” 2018 3rd International Conference on Computer Science and Engineering (UBMK), pp. 625–629, Sep. 2018, doi: 10.1109/ubmk.2018.8566449.

[21] L. Nanni, S. Brahnam, A. Lumini, and G. Maguolo, “Animal sound classification using dissimilarity spaces,” Applied Sciences, vol. 10, no. 23, p. 8578, Nov. 2020, doi: 10.3390/app10238578.

[22] AiProff, "Wildlife Intrusion Detection and Prevention in Farmlands using AI," AiProff, 2022.

[23] S. K. Tiwari, M. A. Qamar, and A. R. Patil, "A Deep Learning-Based Model for Detection and Tracking of Stray Animals in the Fields," Agriculture Journal, vol. 12, no. 3, pp. 56-63, 2023.

[24] R. Das, M. Ghosh, and S. K. R. Ahamed, "Smart Agriculture Land Crop Protection Intrusion Detection Using Image Processing," E3S Web of Conferences, vol. 360, p. 04006, 2023.

[25] S. R. Joshi, "Image Processing-Based Animal Intrusion Detection System in Agricultural Field Using Deep Learning," IEEE Xplore, vol. 10, no. 3, pp. 45-56, 2023

[26] S. L. K. Patel, "Animal Intrusion Detection System Using Mask RCNN," AIP Conference Proceedings, vol. 3075, no. 1, p. 020067, 2023.

[27] Kumar and P. Chaturvedi, "A Review of Existing Farmland Intrusion Detection Systems," International Journal of Computer Applications, vol. 185, no. 22, pp. 97-104, 2023.

[28] R. S. Kumawat, "A Literature Research Review on Animal Intrusion Detection and Repellent Systems," EAI Endorsed Transactions on IoT, vol. 10, no. 4, pp. 45-51, 2021Kumar, "YOLOv4 Implementation for Real-Time Object Detection on Raspberry Pi," IEEE Xplore, 2021

[29] H. Sharma and R. Singh, "Efficient Object Detection on Edge Devices Using YOLOv4 and Raspberry Pi," IEEE Transactions on Industrial Electronics, vol. 68, no. 7, pp. 6025-6033, 2021.

[30] S. Verma, "Real-Time Object Detection with YOLOv4 on Raspberry Pi Using TensorFlow Lite," IEEE Access, vol. 9, pp. 15231-15238, 2021.

[31] L. J. Patel, "YOLOv4 and its Application in Smart Agriculture on Edge Devices," IEEE Access, vol. 9, pp. 11342-11350, 2021.

[32] D. S. Rathi, "A Review of YOLO-based Models for Object Detection on Low-Powered Devices," IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 11, no. 4, pp. 482-490, 2021.

[33] M. X. Rodriguez, "Real-Time Object Detection on Raspberry Pi using YOLOv4-Tiny," TechCrunch, 2021.

[34] L. H. Zhang, "YOLOv4 on Raspberry Pi for Smart Agriculture Applications," IEEE Xplore, 2020.

[35] H. T. Wang, "Edge Computing with YOLOv4 and Raspberry Pi for Real-Time Farm Monitoring," IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 3, pp. 401-410, 2020.