Application of Improved YOLO-LSTM with Combined MQTT-LoRaWAN for AI Surveillance in Tea Plantations to Prevent Elephant Intrusion

Rabin Kumar Mullick*, Rakesh Kumar Mandal
University of North Bengal, Raja Rammohanpur, Darjeeling, West Bengal, India

***Abstract*—Elephant-human conflict is a growing problem in tea garden areas of Dooars in North Bengal, resulting in massive cost for crops, infrastructures and sometimes human life as well. Each year, these mild-mannered giants destroy crops, destroy fences and even threaten the locals, which raises the costs of repairs and endangering lives. The conventional deterring methods, such as fences, firecrackers, and patrols are mostly ineffective, unsustainable or cruel to the animals. As a way of addressing this predicament, scholars have designed a non-invasive, intelligent surveillance system named HIS-Hexagonal Intelligent Surveillance. HIS integrates state-of-the-art machine learning with artificial intelligence by combining an improved YOLO-LSTM and MQTT-LoRaWAN, which combines the capabilities of distributed-based agents with predictive analytics and hex-grid mapping scheme. HIS is an effective solution for elephant intrusions detection and deterrence for ecological balance. The system sends specific warnings before the elephants can cause havoc when an intrusion is detected. The hex-grid mapping enables the operators to have accurate spatial knowledge and the predictive analytics forecasts the time and location where the elephants could roam. The virtual simulation of the proposed work shows 98% accuracy on designed-custom dataset of elephants. The paper offers a background of the architecture, theoretical framework, algorithm models, and expected benefits of the proposed framework.**

***Keywords*—Tea garden; machine learning; artificial intelligence**

I. Introduction

It is the source of livelihood of millions of workers besides being a key economic activity in some areas of Dooars in the North Bengal region of India. These plantations are however, mostly located in or near the elephant habitats. For this reason, human-elephant conflict in the form of elephant crop raiding is a repeated calamity in these areas. Elephants invade plantations for food or migratory movement and cause economic loss, infrastructural damage and casualties.

Reinforcing mahouts using traditional barriers such as electric fencing and manual surveillance for a number of years has not provided the hoped-for relief from elephant pressure. These techniques suffer from little flexibility, laborious nature, and acting in an ecologically or morally questionable manner. There is an urgent need for a smart, non-invasive and autonomous system for dynamic prediction, monitoring and prevention of elephant intrusions.

This study offers a new hybrid strategy to combine different advanced technologies for better tracking and monitoring of elephant movements. It uses Machine Learning (ML) for real-time prediction of the movement pattern of elephants, making a timely and accurate prediction possible. In addition, it adds Intelligence (SI) to autonomously detect and react to changes in the environment and to animal behavior. In order to guarantee a thorough coverage and an efficient zone-based surveillance, the method also utilizes Hexagonal Surveillance (HS) which allows for an efficient spatial surveillance using a structured hexagon-based grid system.

The organization of the present study is as follows: Section II is the review of the related literature, and Section III is the description of the proposed methodology. Section IV identifies the mathematical model. The algorithm appears in Section V, and other Sections provide virtual simulation results, concluding statements and possible research directions.

II. Related Works

Current means of controlling the elephant movement are based on traditional ways, lacking significant limitations. A YOLO-LSTM framework that builds upon the benefit of rapid spatial detection and time modeling and thus delivers very high accuracy in human action recognition in video sequences [1]. Another YOLO-LSTM-based motorway surveillance model, capable of deriving real-time traffic features and forecasting congestion to control effective use of emergency lanes [2, 3]. Rivera-Acosta et al. create the American Sign Language alphabet translator in real time, which puts together a YOLO detector, LSTM module, and a developed spelling correction system [4, 5]. A comparative study of communication protocols in the IoT in CoAP, MQTT, and HTTP in their architecture, performance, and application-suitability [6]. A module-oriented over-the-air firmware update system was proposed on LoRaWAN, which allows flexible, efficient, and reliable distribution of software parts to IoT devices with limited capabilities [7]. Although electric fences are expensive to install and maintain and may be very detrimental to wildlife [8]. Noise-based deterrents like firecrackers are also likely to become ineffective over time as the elephants soon become accustomed to them [9, 10]. Human patrolling though useful is very laborious and unreliable as it is not always available [11].

The recent research indicates that there is the potential of transforming the way machine learning is used to complement various next-generation surveillance systems, such as grid-based approach to guard crop [12]. Cloud-based detection is another research [13]. Among others, one study identifies

*Corresponding author.

the necessity of a communication-aware coverage strategy due to the need to adapt to dynamic environments through monitoring using UAV [14, 15]. Moreover, distributed machine learning has been known to increase capabilities. As a result, as mentioned in a recent open-access archive preprint, UAVs have the capability of providing robust and intelligent surveillance and communication, even in the presence of failures and/or varying conditions through distributed learning [16]. Another study also reported that a hybrid AI with deep reinforcement learning architecture can help in efficient surveillance that increases scalability and autonomy [17]. Collectively, these innovations demonstrate the transformative power of machine learning in driving the frontiers of intelligence for modern surveillance systems.

Swarm-based surveillance has seen a significant advancement in distributed monitoring and intelligent decision making. An instance is the project Eye that offers an integrated platform of intelligence, emergent behavior, and computer graphics. This system permits trying new image processing methods in order to enhance the decision making abilities. It is able to process and capture imagery in parallel via highly developed collaboratively functioning visual units able to view, capture and process imagery simultaneously to create information of great value in real time. Remarkably, the platform is used to provide autonomous and distributed reconfiguration to the UAV formations according to the situational changes in tactical domain, thus exhibiting flexibility as well as adaptability to changing scenarios [18].

Likewise, the DIANA structure solves several important problems in the safety of the motion and effectiveness of operations during -based missions. It presents a new parallel mission planning algorithm that combines evolutionary computation, machine learning, and deterministic solutions to guarantee robust and reliable coordination [19].

Collectively, these innovations demonstrate the potential of networked smart cameras and autonomous agents with machine learning capabilities to serve as intelligent and adaptive components in the advanced surveillance architecture.

Emerging new technologies for wildlife monitoring such as camera traps combined with AI-based image-recognition based on animal tracking, and smart cameras for unending vigil have come into existence but there are serious constraints like shorter battery life, being prone to the environment etc. However, as promising as these technologies are, they are often deployed in silos and are rarely embedded in a convergent methodology or system, almost never with enough real-time agility to adapt for speedy, effective wildlife management.

III. METHODOLOGY

The proposed HIS (Hexagonal Intelligent Surveillance) System — a fully integrated, intelligent framework designed to enhance elephant monitoring and deterrence by combining advanced computational methods and spatial strategies. The schematic overview of the system is shown in Fig. 1. The proposed methodology is shown in Fig. 2. The proposed hexagonal surveillance system is pictured in Fig. 3. The work flow progress of the proposed system is described in Fig. 4. It shows that the proposed system will help to focus the accuracy and speed of the object detection by combining YOLOv8 with Long Short-Term Memory (LSTM) and using MQTT and LoRaWAN to exchange data. The process has three obvious steps, the first step is to deploy a cloud environment to host the computation, the second step is to carefully label and annotate a separate dataset (named elephant) and turn the raw images to a fully ready-to-train model, and, lastly, the combined YOLOv8-LSTM-MQTT-LoRaWAN model is deployed on that dataset, trained, validated and tested its performance to make sure that the improvements are actual. The system is built upon the following core components:

1) Hexagonal / Hex-Grid Area Coverage:

a) Purpose: The surveillance area is divided into hexagons for efficient, non-overlapping coverage.

b) Deployment: Smart cameras are placed at or near each hexagon vertex, ensuring comprehensive monitoring [20].

2) intelligence - Smart Cameras (Agents):

A smart camera can act as an autonomous agent within a surveillance system, using collaborative intelligence to enhance monitoring and situational awareness. Research shows that in distributed surveillance, s function as interconnected units that coordinate flexibly to perform tasks. Smart cameras, as advanced visual agents, capture and process data in parallel, providing valuable insights for collective decision-making [21]. In multi-camera systems, coordinated spatial alignment enables cameras to adjust their positions dynamically for optimal coverage, functioning as intelligent nodes within a broader network [22].

a) Role: Each smart camera acts as an autonomous or semi-autonomous node, capable of local processing and peer-to-peer communication.

b) Collaboration: Cameras share detection results, adaptively adjust focus or coverage as needed [23].

c) Examples: IP (Intenet Protocol) cameras with onboard AI(Artificial Intelligence) (e.g., Reolink Duo 3, TP-Link Tapo C230, or custom smart cameras with edge processing).

3) Onboard Edge Processing (YOLOv8+LSTM): Hybrid YOLOv8-LSTM.

To process this complex data, the system employs a hybrid neural network combining two powerful deep learning models:

a) YOLOv8: The YOLOv8 component plays a key role in spatial feature extraction within the system. It processes image data to identify critical features such as the shape of elephants, movement patterns, and distinct auditory signals. By recognizing these spatial patterns, the YOLOv8 aids in accurately determining the position and behavior of elephants in specific zones, enhancing real-time monitoring and response capabilities.

b) LSTM – Long Short-Term Memory Network: The Long Short-Term Memory (LSTM) network component is responsible for recognizing temporal patterns in elephant behavior. By analyzing sequential data—such as movement trajectories over hours or days—the LSTM learns how elephant activity evolves over time. Its memory units allow it to capture long-term dependencies, enabling the identification of preferred migration routes or seasonal movement trends. This temporal

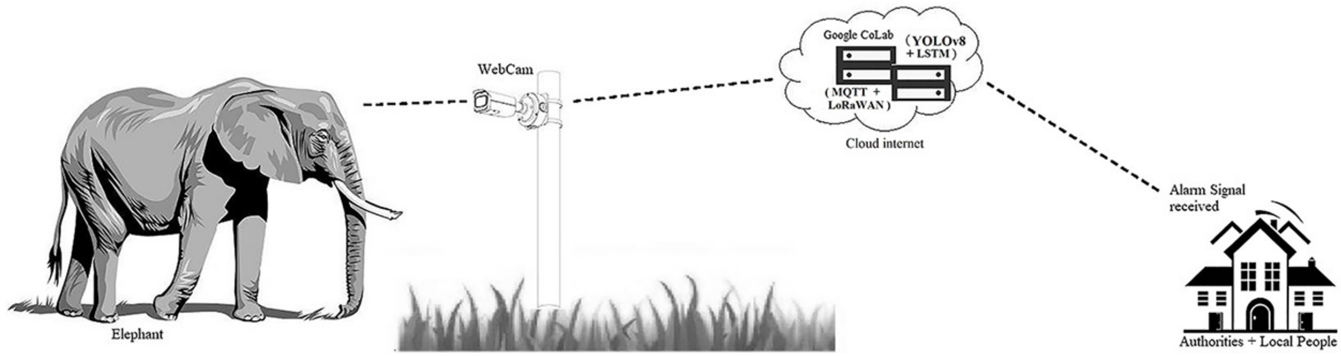

Fig. 1. Schematic diagram overview of the system.

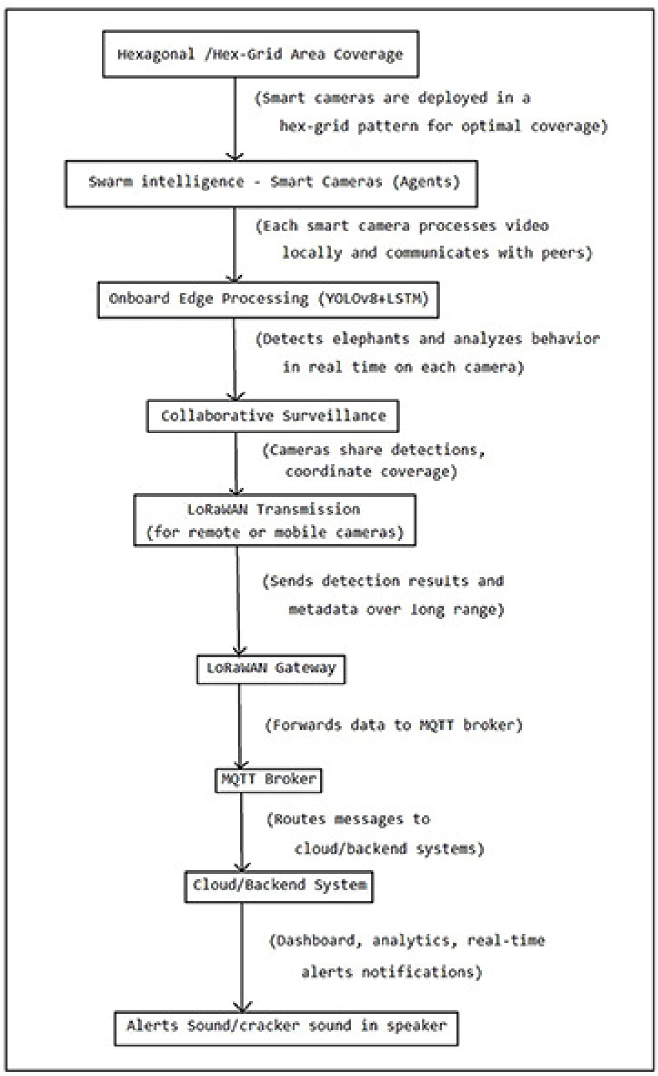

Fig. 2. System architecture: Step-by-step explanation.

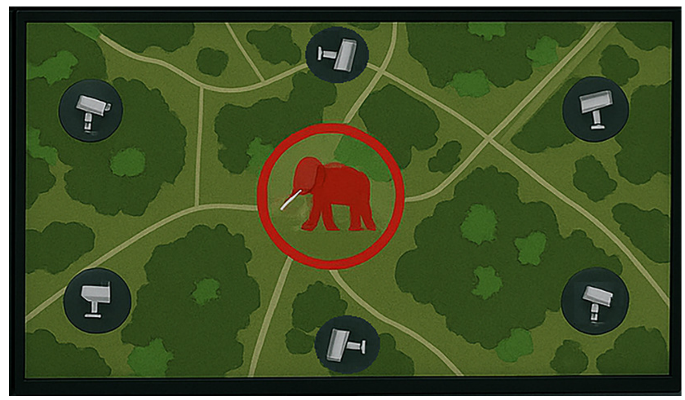

Fig. 3. Hexagonal surveillance system.

insight is crucial for accurately forecasting future elephant movements and enhancing proactive monitoring strategies.

Together, the YOLOv8 detects objects in each frame, while the LSTM models temporal patterns in detection data—enabling the system to predict where elephants are likely to move next, that is, the output it predicts are actions, anomalies, or behaviors.

c) Function: Each camera runs YOLOv8 for real-time elephant detection and LSTM for temporal behavior analysis.

d) Output: Generates detection results (e.g., "elephant detected," "anomaly detected") and metadata.

4) Collaborative Surveillance:

a) Function: Cameras communicate with each other to share detection information, optimize tracking, and ensure continuous coverage.

b) Hand-off: When an object moves from one camera's field of view to another, tracking is seamlessly transferred [24].

5) LoRaWAN Transmission (for remote/mobile cameras):

a) Function: For cameras in remote or hard-to-reach locations, detection results and metadata are transmitted wirelessly via LoRaWAN to a central gateway.

b) Advantage: Long-range, low-power communication suitable for large areas.

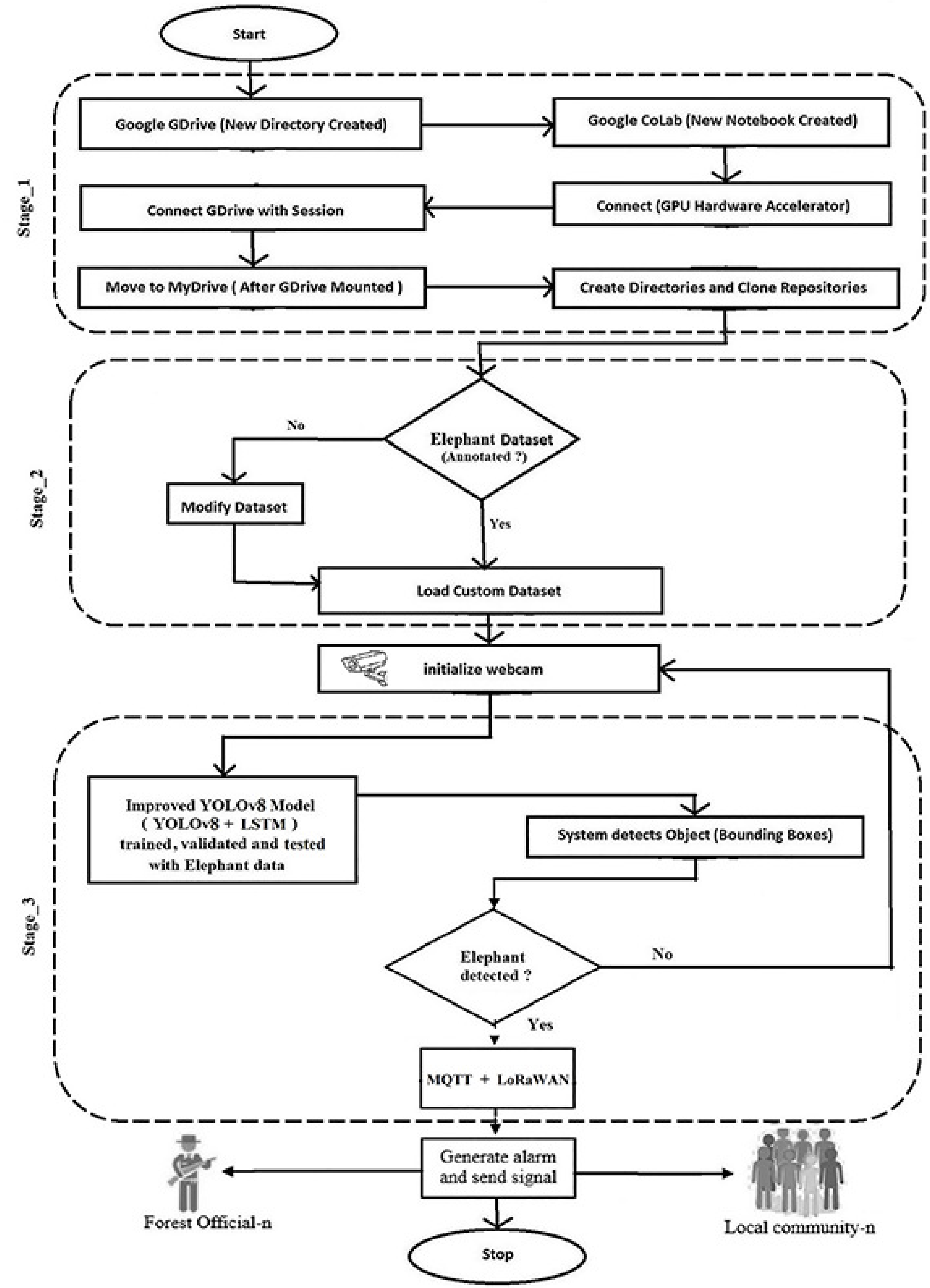

Fig. 4. Work flow progress of the system.

6) LoRaWAN Gateway:

a) Role: Receives data from multiple cameras and forwards it to an MQTT broker.

7) MQTT Broker:

a) Role: Efficiently routes messages to cloud/backend systems for further processing and visualization [25].

8) Backend System:

a) Function: Aggregates data, provides dashboards, analytics, and real-time alerts.

b) User Interface: Authorities or rangers receive notifications and can monitor the area in real time.

9) Alerts sound/ cracker sound in speaker:

a) Purpose: Immediate alerts for elephant detection and deter, enabling rapid response.

10) Key Features:

a) Intelligence: Smart cameras collaborate for adaptive, robust surveillance.

b) Edge AI: YOLOv8+LSTM enable real-time, on-device detection and behavior analysis.

c) Hexagonal Grid: Ensures efficient area coverage and minimal blind spots[26].

d) LoRaWAN & MQTT: Support reliable, scalable communication for remote monitoring [27].

e) Cloud Integration: Provides analytics, dashboards, and real-time alerts.

IV. Mathematical Model for Surveillance System

A. Notations

a) n: Number of smart camera agents in the system.

b) t: Time step (frame index).

c) $C_i(t)$: State of agent i at time t, including its position, battery, and current task.

d) $I_i(t)$: Video frame captured by agent i at time t.

e) $D_i(t)$: Set of detections from agent i at time t, where each detection is a tuple (x, y, w, h, c, p) representing bounding box coordinates, class, and confidence.

f) $\varepsilon_i(t)$: Set of elephant detections (filtered from $D_i(t)$ for class "elephant").

g) $\mathcal{S}_e^{(i)}(t)$: Sequence of features (bounding boxes, confidences) for elephant e tracked by agent i across recent frames.

h) $\text{LSTM}(\cdot)$: LSTM model function.

i) $\mathcal{P}_e^{(i)}(t)$: Temporal prediction (e.g., movement, behavior) for elephant e by agent i at time t.

j) $M_i(t)$: MQTT message published by agent i at time t, containing local predictions and alerts.

k) $G(t)$: Gateway/edge node aggregating MQTT messages from all agents.

l) $M(t)$: Set of all MQTT messages received by the gateway at time t.

m) $A(t)$: Aggregated alert or summary sent via LoRaWAN to control center at time t.

B. Object Detection (YOLOv8)

For each agent i and time t, detections are given by Eq. (1):

$$D_i(t) = \text{YOLOv8}(I_i(t)) \tag{1}$$

Filter for elephant detections, shown in Eq. (2):

$$\varepsilon_i(t) = \{(x, y, w, h, c, p) \in D_i(t) \mid c = \text{"elephant"}\} \tag{2}$$

C. Feature Preparation and Tracking

For each elephant e detected by agent i:

- Track elephant identity across frames (e.g., using a tracking algorithm).
- Build feature sequence for each tracked elephant given in Eq. (3):

$$\mathcal{S}_e^{(i)}(t) = [(x, y, w, h, p)_{t-k}, \ldots, (x, y, w, h, p)_t] \tag{3}$$

where k is the sequence length.

D. Temporal Analysis (LSTM)

The Eq. (4) shows the temporal prediction using LSTM:

$$\mathcal{P}_e^{(i)}(t) = \text{LSTM}(\mathcal{S}_e^{(i)}(t)) \tag{4}$$

E. Coordination (MQTT)

Each agent publishes its predictions and alerts, as shown in Eq. (5):

$$M_i(t) = \{(\mathcal{P}_e^{(i)}(t), \text{agent\_id}, \text{timestamp}) \mid e \in \varepsilon_i(t)\} \tag{5}$$

Gateway collects messages from all agents, shown in Eq. (6):

$$M(t) = \bigcup_i M_i(t) \tag{6}$$

F. Data Aggregation and Transmission (LoRaWAN)

Gateway aggregates and compresses messages for transmission shown in Eq. (7):

$$A(t) = \text{Aggregate}(M(t)) \tag{7}$$

G. Summary Equation

Overall surveillance system model is shown in Eq. (8):

$$A(t) = \text{Aggregate}\Big(\{\text{LSTM}(\text{Track}(\text{YOLOv8}(I_i(t))))\}\Big) \tag{8}$$

V. Proposed Algorithm

The proposed work has been described in the below Algorithm 1:

Algorithm 1 Hexagonal Surveillance System with YOLOv8, LSTM, MQTT and LoRaWAN integration

Initialize
Deploy η smart camera agents across the surveillance area
Set up MQTT broker and LoRaWAN gateway
Initialize YOLOv8 and LSTM models on each agent
Define sequence length κ for LSTM input
Initialize tracking for elephant identity maintenance
Compute
For each agent i, capture initial video frame $I_i(0)$
while surveillance is active **do**
  for every agent i **do**
    Capture current video frame $I_i(t)$
    Detect objects using YOLOv8: $D_i(t) = \text{YOLOv8}(I_i(t))$
    Filter for elephant detections:

$$\varepsilon_i(t) = \{(x, y, w, h, c, p) \in D_i(t) \mid c = \text{“elephant”}\}$$

    Track each elephant $e \in \varepsilon_i(t)$ across frames
    Build feature sequence for each tracked elephant
    Compute temporal prediction: $\mathcal{P}_e^{(i)}(t) = \text{LSTM}(\cdot)$
    Publish prediction and alert via MQTT:

$$M_i(t) = \{(\mathcal{P}_e^{(i)}(t), \text{agent\_id}, \text{timestamp}) \mid e \in \varepsilon_i(t)\}$$

    if unusual behavior or critical alert **then**
      Search for optimal agent coverage or notify control center
    end if
  end for
  Gateway collects all MQTT messages: $M(t) = \bigcup_i M_i(t)$
  Gateway aggregates and transmits via LoRaWAN: $A(t) = \text{Aggregate}(M(t))$
end while

VI. Virtual Simulation

The proposed work has been performed in Google Colab with designed-custom dataset. The virtual simulation of the proposed methodology has been shown in following sections:

A. Construction of Dataset

In this research, we compiled a dedicated dataset, which we obtained through the web by gathering the images of elephants. The range of poses, lighting circumstances, and backgrounds covered by these photographs is quite extensive, thus allowing the model to learn how to see elephants in many real-life situations. We additionally used a basic Python script to cut still frames out of video footage thus further increasing the data pool. After compilation, the images may be enhanced using a number of deep-learning algorithms to enhance quality and consistency then to train the detection system.

B. Data Annotation

The open-source tool of labeling named as LabelImg redefines the labeling process in such a way that it does not feel like a tedious experience but instead a creative and simplified one. The user can easily outline bounding boxes on parts of an image that are of interest, apply class labels to them, and save that information in a one-step process. Internally, the tool creates a plain-text document where the first integer is the identifier of the class, then there is the coordinates of the centre of the box and the width and height of the box. Since LabelImg can also export annotations in both YOLO and PAScal VOC XML formats, as well as CreateML, the cost of an additional conversion step is eliminated, especially with systems that operate on the YOLO framework. The software also offers easy key board shortcuts to drawing and saving and viewing images and allows remote server annotation. , therefore, regardless of the intended goal, be it detection, segmentation, or classification, the intuitive workflow and pre-defined system of classes provided by LabelImg allows the user to focus on the underlying data, instead of focusing on the underlying mechanics of labeling.

C. Dataset Preparation

The present project created a strong and detailed dataset when the real-world observations were integrated with the synthetic information to improve the predictive model in the real-time applications. The real-world data were obtained by the means of the open-access datasets of the elephant movements.

The models are trained using the custom dataset using 8,608 number of original elephant images and 1,592 number of non-elephant images. Table I has presented the information in the image dataset used to evaluate the models.

TABLE I. Designed-Custom Dataset

Image	Training	Validation	Test	Total
Positive (Elephant)	6,458	2,150	2,150	10,758
Negative (Non_elephant)	1,192	400	400	1,992
Total	7,650	2,550	2,550	12,750

D. Google CoLab

Google is a cloud based, no cost Jupyter notebook bearing the name Google Colab designed to engage a group of coders in a game. It also provides free and immediate access to strong GPUs and TPUs, ideal when training machine-learning models, and also allows to share a notebook with teammates as with a Google Doc. Everything on Google works in Google Drive, and it includes the possibility to open, edit, and share files everywhere. and since all notebooks are already pre-equipped with popular data science libraries such as TensorFlow, PyTorch, and pandas amongst others, one can immediately get started on data science instead of wasting time installing on own environment.

E. Model Training

The hybrid YOLOv8-LSTM architecture, i.e., the spatial feature extraction of YOLOv8 and the temporal sequence modeling of Long Short-Term Memory (LSTM) networks, was trained to create the predictive model. This architecture was chosen to be good at retrieving both spatial patterns (e.g., movement paths over terrain), and temporal dynamics (e.g., repeat entry times, or repeat behavior sequences) related to the activity of elephants. To balance the evaluation, an 60-20-20 (train-validate-test split) was used to divide the dataset in order to allow the model to generalize well to unknown data.

When trained, the model was very impressive and detected the entry patterns of the elephants with a high level of accuracy of about 98 percent. The findings support the strength of the model in practice, and it is highly suitable when applying to the proactive wildlife surveillance and early-warning systems.

VII. Simulated Results

The model evaluation was developed in controlled simulation conditions to determine the effectiveness of the system, responsiveness and efficiency of the system.

The confusion matrix for proposed ML model is shown in Fig. 5.

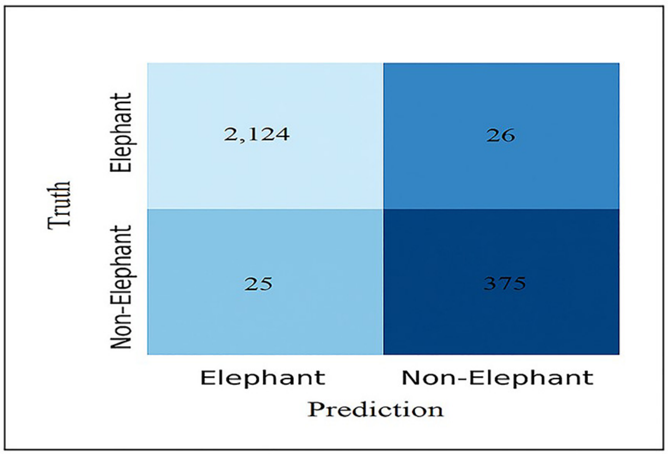

Fig. 5. Confusion matrix for proposed ML model.

The Fig. 6 shows confusion matrix for YOLOv8 model for elephant detection.

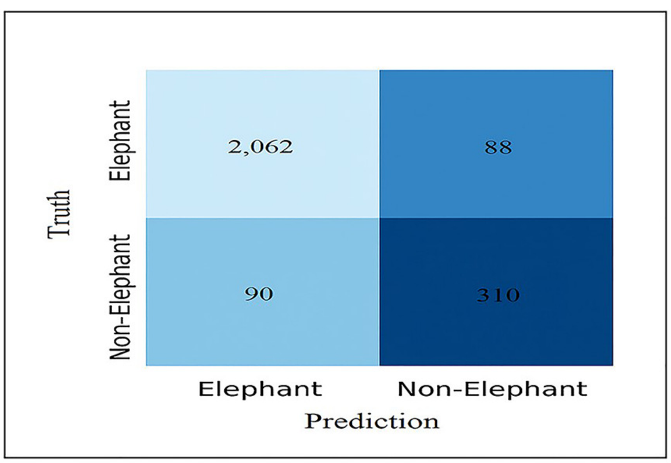

Fig. 6. Confusion matrix for YOLOv8 model for elephant detection.

Table II presents the results Machine Learning(ML) of proposed model.

The Fig. 7 shows the elephant detected by the proposed system.

The prediction model had a high accuracy of 98.0 percent in detecting the patterns of elephant movement and low false positive of 1.0 percent that guaranteed reliability of the detection with minimal false alarms.

TABLE II. ML Results for Proposed Model

Parameter (/ Measure)	Proposed ML	YOLOv8
Total predicted	2,550	2,550
True positive	2,124	2,062
False positive	26	88
False negative	25	90
True negative	375	310
Accuracy	0.98	0.93

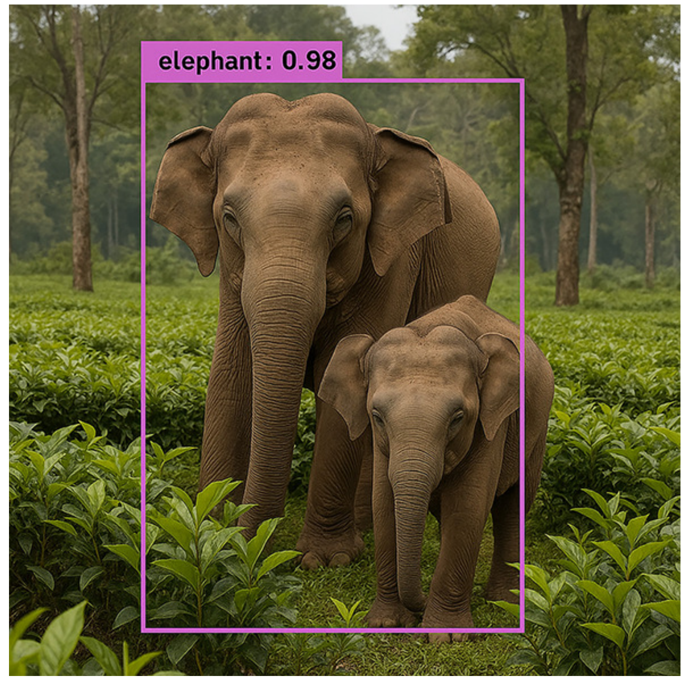

Fig. 7. Elephant detected.

VIII. Discussion

With MQTT on top of LoRaWAN, it is possible to make AI-based surveillance systems both practical and large-scale. LoRaWAN oxygen incapacity has some sweeping advantages of IoT links to hundreds of inactive, long-range sensors in a town, villa, or off-grid locations that maintain their operation with minimal roadwork studies. MQTT then manages the light, latency constrained communications between those edge devices and the cloud, cuts down bandwidth wastage and allows the network to scale to any topography; be it a smart village or a microgrid [28]. It is a combination of feeds live to YOLO-LSTM models providing video and sensor data in real time reliably even when in dense forests or during busy city times, therefore object detection and sequence analysis remains keen in high-traffic surges. Collectively, MQTT and LoRaWAN reduce power consumption and data rates that enable cameras and sensors detached in the distant areas to extend their lifespan on a limited battery and yet still provide the real time information that current AI surveillance relies upon [29, 30].

Table III presents the key advantages of proposed model.

TABLE III. SUMMARY TABLE: KEY ADVANTAGES

Parameter (/ Feature)	LoRaWAN Contribution	MQTT Contribution	Combined with YOLO-LSTM for Benefiting AI Surveillance
Coverage	Long-range	Not fixed	Flexible deployment
Power Consumption	Minimal	Reduces unnecessary transfer	Sustainable, low-maintenance operation
Scalability	Supports many devices	Flexible brokering	Easily expandable networks
Latency	Low (with MQTT integration)	Efficient message handling	Real-time data for AI models
Data Congestion	Avoids with hybrid protocol	Manages high-frequency data	Reliable, up-to-date AI inference

Overall, the integration of MQTT and LoRaWAN with an YOLO-LSTM model to create AI surveillance systems has major scalability, power consumption, low-latency communications, environmental flexibility benefits which are essential to advanced and efficient real-time surveillance systems.

IX. FUTURE PROPOSAL

The current simulation-based system shows a satisfactory initial result, which can be further developed in future work for a wider scope of practical application and technology robustness. Finally, field trials should be executed in cooperation with the forest departments and conservation authorities to ensure the performance of the system in real conditions and to improve its behavior under natural environmental conditions. These trials are expected to give important feedback regarding agent coordination, detection accuracy, and system robustness in various terrain types. Additionally, blockchain technology is suggested to be implemented to provide transparency and security of the data, enabling tamper-proof and decentralized data sharing between, for example, tea estate owners, conservationists, or local administration. This would provide confidence and accountability of system warning and movement records. In addition, thermal imaging sensors will be integrated for intelligences through enhanced detection capabilities particularly at night, when the elephant's skin color is barely noticeable by agents, irrespective of the lighting conditions. Additionally, the capabilities to create mobile applications that can provide localized parties (planting workers and local communities) with real-time alerts will allow localized response initiatives, enabling even more human-autonomous system collaboration. Together, these improvements will bring the system closer to full-scale deployment, and bring about a robust, ethical, community-driven elephant conflict mitigation solution.

X. CONCLUSION

This paper introduces a new, holistic and interdisciplinary way of looking at the increasingly challenging issue of elephant-human conflict, especially in tea plantation areas. The proposed autonomous Hexagonal Intelligent Surveillance (HIS) entails the non-separate, combined application of machine learning-based prediction models, intelligence, and a hexagonal spatial grid to constitute a cohesive, responsive, and ethically-grounded system. The virtual simulation of the proposed work shows 98% accuracy on designed-custom dataset of elephants. In addition, in order to reduce the need for human intervention, the system detects potential elephant intrusions ahead of time and provides coordinated adjuvant decentralized reaction through autonomous agents in a non-lethal and animal safe way. The hexagonal cell topology improves the coverage and flexibility of the system by enabling it to adapt to changing levels of risk over large geographical areas. Apart from its scalability, this combination of technologies puts both the safety of humans and the conservation of wildlife at the forefront. With further field testing and real-world implementation, the HIS framework has the potential of revolutionizing current strategies for wildlife conflict mitigation by providing a smarter, more efficient, and more humane solution for preserving livelihood and ecosystems on agricultural frontiers worldwide.

ACKNOWLEDGMENTS

The authors sincerely thank the anonymous reviewers for their valuable comments and constructive suggestions, which significantly improved the quality of this paper. Their insightful feedback is greatly appreciated.

DISCLOSURE OF INTERESTS

The authors have no competing interests to declare that are relevant to the content of this article.

REFERENCES

[1] Elnady, Mahmoud, and Hossam E. Abdelmunim. "A novel YOLO LSTM approach for enhanced human action recognition in video sequences." Scientific Reports 15.1 (2025): 17036.

[2] Tong, Xing, et al. "Efficient Utilization Method of Motorway Lanes Based on YOLO-LSTM Model." Sensors 25.21 (2025): 6699.

[3] Reis, Dillon, et al. "Real-time flying object detection with YOLOv8." arXiv preprint arXiv:2305.09972 (2023).

[4] Rivera-Acosta, Miguel, et al. "Spelling correction real-time american sign language alphabet translation system based on yolo network and LSTM." Electronics 10.9 (2021): 1035.

[5] Lu, Yongyi, Cewu Lu, and Chi-Keung Tang. "Online video object detection using association LSTM." Proceedings of the IEEE international conference on computer vision. 2017.

[6] Ruchi, Lalhriatpuii, and Vikas Wasson. "Comprehensive Exploration of IoT Communication Protocol: CoAP, MQTT, HTTP." Machine Learning Algorithms: First International Conference, ICMLA 2024, Himachal Pradesh, India, February 23–24, 2024, Proceedings. Vol. 2238. Springer Nature, 2024.

[7] Nguyen, Huy Dat, Nicolas Le Sommer, and Yves Mahéo. "Over-the-Air Firmware Update in LoRaWAN Networks: A New Module-based Approach." Procedia Computer Science 241 (2024): 154-161.

[8] Kioko, John, et al. "The performance of electric fences as elephant barriers in Amboseli, Kenya." South African Journal of Wildlife Research-24-month delayed open access 38.1 (2008): 52-58.

[9] Sivakumar, V., et al. "Decibel level of firecrackers and its possible impact on the hearing of marauding elephants in Sri Lanka." Int. J. Sci. Environ. Technol. 2.4 (2013): 593-600.

[10] King, Lucy E., Iain Douglas-Hamilton, and Fritz Vollrath. "African elephants run from the sound of disturbed bees." Current Biology 17.19 (2007): R832-R833.

[11] Montgomery, Robert A., et al. "The efficacy of interventions to protect crops from raiding elephants." Ambio 51.3 (2022): 716-727.

[12] Mullick, Rabin Kumar, and Rakesh Kumar Mandal. "A Proposed Grid-Based Elephant Detection Model Using Artificial Intelligence (AI) to Prevent Crop Damage in Farming Fields." Doctoral Symposium on Intelligence Enabled Research. Singapore: Springer Nature Singapore, 2023.

[13] Mullick, Rabin Kumar, and Rakesh Kumar Mandal. "An Efficient Elephant Detection Strategy using Visual Attention Network (VAN) in Custom Dataset improved YOLOv7 Model." (2025).

[14] Tahir, Anam, et al. " s of unmanned aerial vehicles—A survey." Journal of Industrial Information Integration 16 (2019): 100106.

[15] Xu, Chengtao, et al. "Communication aware UAV surveillance based on hierarchical architecture." Drones 5.2 (2021): 33.

[16] Ding, Yahao, et al. "Distributed machine learning for uav s: Computing, sensing, and semantics." IEEE Internet of Things Journal 11.5 (2023): 7447-7473.

[17] Arranz, Raul, et al. "Application of deep reinforcement learning to UAV ing for ground surveillance." Sensors 23.21 (2023): 8766.

[18] Khan, Faisal, et al. "Eye: A Distributed Autonomous Surveillance System."International Journal of Advanced Computer Science and Applications (IJACSA), 9.12 (2018): 591-600.

[19] Majd, Amin. "DIANA: Distributed and Safe Autonomous Navigation for a of Autonomous Vehicles." (2021).

[20] A. Kumar, S. Patel, R. Singh. "Elephant recognition using deep learning: A comparative study." Data in Brief, 2024. https://www.sciencedirect.com/science/article/abs/pii/S2352710223014377.

[21] Tian, Wei, et al. "A Review of Smart Camera Sensor Placement in Construction." Buildings 14.12 (2024): 3930.

[22] Li, Congcong, et al. "Multi-camera joint spatial self-organization for intelligent interconnection surveillance." Engineering Applications of Artificial Intelligence 107 (2022): 104533.

[23] B. Rinner. "Distributed Smart Cameras." Technische Universität Graz, 2008. https://www.ece.rice.edu/ av21/Documents/pre2011/Distributed%20 Smart%20Cameras.pdf.

[24] T. Menteş, H. Artuner, C. Baransel. "A Distributed Surveillance Camera Network Architecture for Public Security Applications in Metropolitan Area." Hacettepe University, 2010.

[25] Amazon Web Services. "Design considerations for cost-effective video surveillance platforms with AWS IoT for Smart Homes." AWS IoT Blog, 2023. https://aws.amazon.com/blogs/iot/design-considerations-for-cost-effective-video-surveillance-platforms-with-aws-iot-for-smart-homes.

[26] An, Zehua, Xiaoping Rui, and Chaojie Gao. "Improved A* navigation path-planning algorithm based on hexagonal grid." ISPRS International Journal of Geo-Information 13.5 (2024): 166.

[27] L. Valentín, S.A. Serrano, R. Oves García, A. Andrade, M.A. Palacios-Alonso, L.E. Sucar. "A Cloud-Based Architecture for Smart Video Surveillance." The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-4/W3, 2017. https://d-nb.info/1143989678/34.

[28] Shuaib M., Alam S. (2025). A secure and energy-efficient IoT-blockchain framework for decentralized renewable energy trading. Sustainable Computing: Informatics and Systems, 48. https://doi.org/10.1016/j.suscom.2025.101252.

[29] Changazi S., Bakhshi A., Yousaf M., Mohsin S., Akber S., Abazeed M., Ali M. (2024). Optimization of network topology robustness in IoTs: A systematic review. Computer Networks, 250. https://doi.org/10.1016/j.comnet.2024.110568.

[30] Hadji O., Maimour M., Benyahia A., Kadri O., Rondeau E. (2025). EcoWatch: Region of interest-based multi-quantization resource-efficient framework for migratory bird surveillance using wireless sensor networks and environmental context awareness. Computers and Electrical Engineering, 123. https://doi.org/10.1016/j.compeleceng.2025.110076.