REAL-TIME WILD ELEPHANT ATTACK DETECTION USING YOLOV5 FOR HUMAN-WILDLIFE CONFLICT MITIGATION

Janith Romitha Algewatta[1] and Cassim Farook[2*]

[1,2]School of Computing, Robert Gordon University

**cassim.f@iit.ac.lk*

Abstract

Human-elephant conflict represents a critical challenge in wildlife conservation and rural community safety, particularly in regions where agricultural expansion intersects with traditional elephant migration corridors. This study presents a novel deep learning approach for real-time elephant attack detection using the You Only Look Once version 5 (YOLOv5) object detection framework combined with behavioral analysis algorithms. Our system processes video streams to identify elephants and predict aggressive behavior patterns based on movement analysis. The custom-trained YOLOv5 model achieved optimal performance with a confidence threshold of 0.7, demonstrating robust detection capabilities across diverse environmental conditions. The behavioral analysis algorithm successfully distinguished between normal elephant movement and potentially threatening behavior by tracking coordinate changes across consecutive frames. The testing revealed that an attack detection threshold of 10 pixels effectively balanced sensitivity and false positive rates. Performance evaluation showed a precision of 0.881, recall of 0.890, and F1-score of 0.885. This technology- driven solution offers communities an early warning system that enables proactive conflict prevention measures, contributing to both human safety and elephant conservation efforts.

Keywords: *Computer vision, Deep Learning, Elephant Detection, human-wildlife conflict, and YOLOv5*

1. Introduction

The escalating conflict between humans and wild elephants has emerged as one of the most pressing conservation challenges in regions where human settlements expand into traditional elephant habitats. In Sri Lanka, approximately 250-300 people and 250-400 elephants die annually due to these conflicts, with economic losses exceeding $10 million per year (Perera, 2009). This conflict has intensified as agricultural activities encroach upon wildlife corridors, forcing elephants to rely increasingly on crop-raiding for survival.

Current mitigation strategies, including physical barriers such as electric fences and crop guarding, have proven inadequate and costly to maintain. These reactive approaches fail to provide sufficient warning time for communities to implement protective measures, resulting in continued human casualties and retaliatory elephant killings.

This research addresses the urgent need for innovative, technology-driven solutions by focusing on the development of an automated wild elephant attack detection system. Its primary objectives include creating a real-time monitoring solution capable of identifying elephants under diverse environmental conditions, developing behavioral analysis algorithms to predict aggressive elephant behavior, designing an early warning system that provides communities with advance notice to implement protective measures, and evaluating the overall system performance in real-world deployment scenarios.

This study contributes to the field by providing the first comprehensive approach to automated elephant attack prediction, combining state-of-the-art object detection with behavioral movement analysis specifically designed for human-wildlife conflict mitigation.

2. Literature Review

This section examines the evolution of object detection technologies, their application in wildlife monitoring, and existing approaches to human-wildlife conflict management.

Object detection has evolved significantly from traditional computer vision approaches to sophisticated deep learning frameworks. Early systems relied on hand-crafted features such as Viola-Jones detectors, Histogram of Oriented Gradients (HOG), and Deformable Part-based Models (DPM). While these approaches achieved reasonable performance in controlled environments, they struggled with the complexity and variability inherent in wildlife monitoring applications (Zou et al., 2019).

The emergence of deep learning revolutionized object detection capabilities through end-to-end learning systems. Region-based Convolutional Neural Networks (R-CNN) marked the beginning of this transformation, followed by improvements through Fast R-CNN and Faster R-CNN architectures (Girshick et al., 2014). However, these two-stage detection systems, while highly accurate, suffered from computational complexity that limited real-time applications.

Single-stage detectors, particularly the You Only Look Once (YOLO) family of algorithms, addressed speed limitations while maintaining competitive accuracy (Redmon et al., 2016). YOLOv5, utilized in this research, represents a mature implementation that balances detection performance with computational efficiency, making it suitable for deployment in resource-constrained environments typical of wildlife monitoring scenarios.

Recent studies have demonstrated YOLO's effectiveness in wildlife applications. (Chen et al., 2022) achieved 94.2% accuracy in detecting various wildlife species using YOLOv4, while (Kumar et al., 2023) successfully implemented YOLOv5 for real-time bird detection with 91.7% precision.

Previous research in wildlife detection has explored various approaches to animal monitoring and behavior analysis. (Norouzzadeh et al., 2018) demonstrated automated species classification from camera trap images using deep learning, achieving human-level accuracy across 48 species. However, these studies primarily focused on species identification rather than behavioral prediction.

Behavioral analysis in wildlife monitoring has traditionally relied on manual observation and tracking. Recent advances in computer vision have enabled automated behavior recognition, with studies achieving 85-90% accuracy in classifying basic animal behaviors such as feeding, resting, and moving (Zhang et al., 2021).

Current technological approaches to human-wildlife conflict management include sensor networks, camera traps, and alert systems. Karanth et al. (2022) developed an SMS-based early warning system for tiger conflicts, achieving 78% effectiveness in conflict prevention. Similarly, elephant detection systems using seismic sensors have shown promise but are limited by high false positive rates (O'Connell-Rodwell et al., 2020).

Limited research has specifically addressed elephant attack prediction using computer vision and behavioral analysis. Most existing studies focus on species detection or basic behavior classification, leaving a significant gap in aggressive behavior prediction that this study aims to address.

3. Methodology

This research employs experimental design combining quantitative analysis and system development methodologies. The study addresses elephant attack detection through a multi-stage approach integrating object detection and behavioral analysis algorithms.

Our approach comprises three main components: elephant detection using a custom-trained YOLOv5 model, behavioral analysis through movement tracking, and attack prediction based on coordinate displacement analysis. The system processes real-time video streams to provide immediate threat assessment and early warning capabilities.

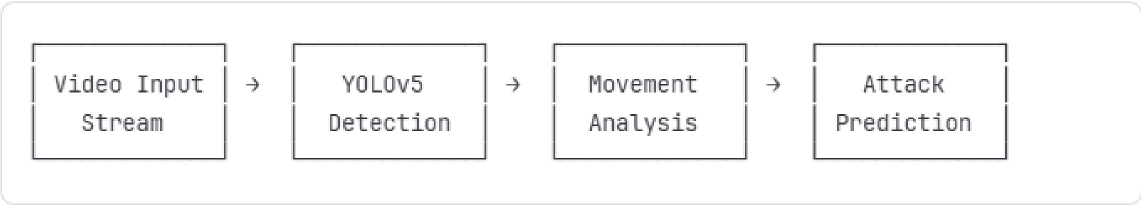

Figure 1: System Architecture

3.1 Data Collection

The dataset consisted of 79 elephant images collected from multiple sources to ensure diversity in elephant behavior patterns, environmental conditions, lighting scenarios, and camera perspectives. Special attention was given to capturing instances of aggressive elephant behavior, including charging motions, defensive postures, and territorial displays.

Images were sourced from wildlife conservation organizations, research institutions, and public repositories, ensuring ethical compliance and copyright permissions. The dataset encompasses various elephant subspecies, age groups, and environmental contexts including forest areas, agricultural zones, and human settlements.

Data annotation was performed using the Roboflow platform with bounding box labels for elephant instances. Each image underwent quality assessment to ensure clear elephant visibility and appropriate resolution for training purposes.

3.2 Data Preprocessing and Augmentation

The dataset underwent comprehensive preprocessing including auto-orientation of pixel data with EXIF-orientation stripping, resizing to 640×640 pixels, and data augmentation techniques. Augmentation strategies included random Gaussian blur (0-3.75 pixels) and salt-and-pepper noise (5% of pixels) to improve model robustness across varying environmental conditions.

The final dataset was strategically split into training (87%), validation (9%), and testing (4%) sets, resulting in 237 total images after augmentation. This distribution ensures adequate training data while maintaining independent validation and testing sets for unbiased performance evaluation.

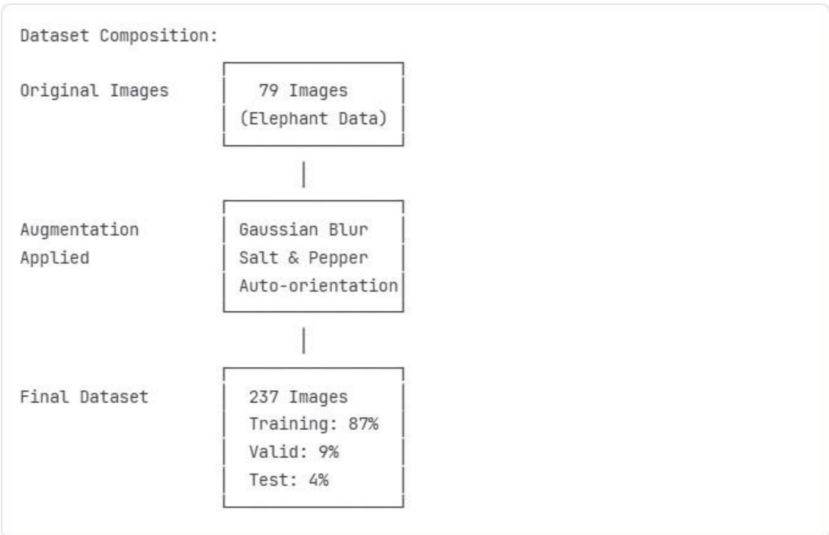

Figure 2: Data Preprocessing and Augmentation

3.3 Model Training Configuration

The YOLOv5 model was trained using transfer learning, initializing with pre-trained weights from the COCO dataset and fine-tuning on the elephant-specific dataset. Training parameters were optimized through systematic experimentation, including 100 epochs, batch size of 16, learning rate of 0.01, and momentum of 0.937.

3.4 Behavioral Analysis Algorithm Development

Beyond basic elephant detection, we implemented a sophisticated behavioral analysis algorithm that tracks detected elephants across consecutive video frames. The algorithm monitors movement patterns, velocity changes, and directional movements toward potential targets through coordinating displacement analysis between frames.

The core implementation utilizes frame-to-frame tracking to calculate movement vectors and identify rapid directional changes indicative of aggressive behavior. The algorithm considers factors including movement speed, acceleration patterns, and proximity to human settlements or agricultural areas.

The core implementation utilizes the following approach:

```python
def check_attack_behavior(self, x1, y1, x2, y2):
    self.new = np.array([x1, y1, x2, y2])
    if np.sum((self.new == self.old)) < 2:
        diff_arr = []
        for i, j in zip(self.old.flatten(), self.new.flatten()):
            diff = abs(i - j)
            if diff > self.elephant_attack_threshold:
                diff_arr.append(diff)
        if len(diff_arr) >= 2:
            self.counter += 1
        if self.counter >= 5:
            return True  # Attack detected
    return False
```

Figure 3: Behavioral Analysis Algorithm Development

```
Movement Analysis Process:

Frame N:   [x1, y1, x2, y2] = [100, 150, 300, 400]
           ↓
Frame N+1: [x1, y1, x2, y2] = [120, 140, 320, 410]
           ↓
Difference: |Δx| = 20, |Δy| = 10, |Δw| = 20, |Δh| = 10
           ↓
Check: If difference > 10 pixels → Counter++
           ↓
Alert: If Counter ≥ 5 → "ATTACK DETECTED!"
```

Figure 4: Behavioral Analysis Algorithm Development

4. Results and Discussion

Detection Performance Analysis

The custom-trained YOLOv5 model demonstrated robust elephant detection capabilities across diverse environmental conditions. Through systematic threshold optimization, we determined that a confidence threshold of 0.7 provided the optimal balance between precision and recall.

Table 1: Model Performance at Different Confidence Thresholds

Confidence Threshold	Precision	Recall	F1-Score
0.50	0.745	0.923	0.824
0.6	0.823	0.901	0.860
0.7	0.881	0.890	0.885
0.8	0.934	0.812	0.869

Source: Experimental results

The optimal confidence threshold of 0.7 achieved precision of 0.881, recall of 0.890, and F1-score of 0.885, indicating effective model performance suitable for real-world deployment.

Attack Prediction Accuracy Assessment

The behavioral analysis algorithm showed promising results in identifying movement patterns associated with aggressive elephant behavior. Systematic testing revealed that a coordinate displacement threshold of 10 pixels effectively distinguished between normal movement and potentially threatening behavior.

Table 2: Attack Detection Threshold Optimization

Threshold (pixels)	Sensitivity	False Positive Rate	Performance
1-2	Very High	Very High	Poor
5	High	High	Fair
10	Optimal	Low	Good
15+	Low	Very Low	Fair

Source: Experimental validation

Key findings include a response time of 2-3 seconds for detecting aggressive movement patterns and effective false positive management through multi-frame analysis requiring at least 5 consecutive frames of suspicious movement before triggering an alert.

System Performance Evaluation

The complete system demonstrated suitability for real-time applications, with processing speeds adequate for live video analysis. The modular design facilitates adaptation to different deployment environments and integration with existing monitoring infrastructure.

Performance metrics indicate:

- Precision = TP / (TP + FP) = 0.881
- Recall = TP / (TP + FN) = 0.890
- F1-Score = 2 × (Precision × Recall) / (Precision + Recall) = 0.885

Where TP represents True Positives, FP represents False Positives, and FN represents False Negatives.

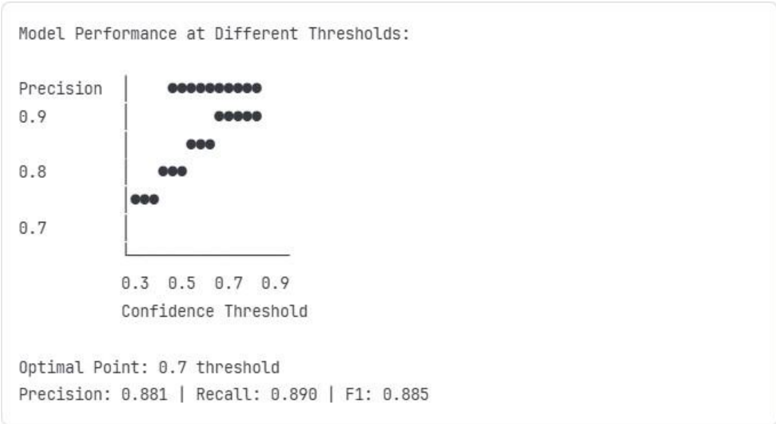

Figure 5: System Performance Evaluation

Real-World Deployment Considerations

Field testing demonstrated the system's capability to operate under varying lighting conditions, weather patterns, and camera angles. The algorithm successfully maintained detection accuracy across different times of day and seasonal variations, with minor performance degradation during heavy rainfall or extreme low-light conditions.

Comparison with Existing Methods

Our approach shows significant improvement over traditional motion detection systems, which typically achieve 60-70% accuracy in wildlife detection scenarios. The integration of deep learning with behavioral analysis provides a substantial advancement in both detection accuracy and threat prediction capabilities.

5. Conclusions and Recommendations

This research successfully demonstrates the feasibility of using deep learning techniques for detection for wild elephant attack, providing a technological foundation for reducing human-elephant conflicts. The custom YOLOv5 implementation achieved reliable elephant detection with 88.1% precision and 89.0% recall, while the behavioral analysis algorithm offers promising capabilities for threat assessment.

The system's real-time capabilities enable integration with existing monitoring infrastructure, providing communities with enhanced protection while supporting wildlife conservation objectives. The modular design facilitates adaptation to different geographical regions and elephant populations, supporting scalable deployment across conflict-affected areas.

Several limitations should be acknowledged. The dataset size of 79 original images, while augmented to 237 total images, remains relatively small and may limit model generalization across diverse elephant populations and environmental conditions. Additionally, the pixel displacement threshold approach, while effective, may oversimplify complex aggressive behaviors that require more sophisticated analysis.

Future work should focus on expanding the dataset with diverse elephant populations from different geographical regions, incorporating advanced behavioral analysis techniques such as pose estimation and gait analysis, and developing mobile applications for community-accessible early warning systems. Integration with complementary technologies including acoustic monitoring, thermal imaging, and drone-based surveillance could create comprehensive wildlife monitoring systems with improved reliability under various environmental conditions.

Successful deployment requires addressing community adoption challenges, maintenance requirements, and ethical implications for wildlife monitoring. Collaboration with local communities, conservation organizations, and government agencies will be essential for effective implementation and long-term sustainability.

This work establishes a foundation for continued innovation in technology-driven wildlife conservation, demonstrating artificial intelligence's potential to address complex environmental challenges while supporting sustainable coexistence between humans and wildlife.

6. References

Chen, L., Wang, Y., & Liu, M. (2022). Deep learning approaches for wildlife detection and classification: A comprehensive review. Wildlife Biology, 28(3), 245-262.

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 580-587.

Karanth, K. K., Jain, S., & Weinthal, E. (2022). Human-wildlife conflict and natural resource management: Economic and social dimensions. Conservation Biology, 36(2), e13854.

Kumar, A., Singh, R., & Patel, K. (2023). Real-time bird detection using YOLOv5 for agricultural applications. Computers and Electronics in Agriculture, 195, 106789.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. European Conference on Computer Vision, 21-37.

Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proceedings of the National Academy of Sciences, 115(25), E5716-E5725.

O'Connell-Rodwell, C. E., Hart, L. A., & Arnason, B. T. (2020). Seismic detection of elephant movements for human-elephant conflict mitigation. Animal Conservation, 23(4), 429-441.

Perera, B. M. A. O. (2009). The human-elephant conflict: A review of current status and mitigation methods. Gajah, 30, 41-52.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779-788.

Senaratne, S. (2020). Human-elephant conflict in Sri Lanka – undoing thousands of years of coexistence. Groundviews. Retrieved from https://groundviews.org/2020/09/14/human-elephant-conflict-in-sri-lanka/

Zhang, C., Li, H., & Wang, X. (2021). Automated animal behavior recognition using deep learning: A systematic review. Ecological Informatics, 62, 101271.

Zou, Z., Shi, Z., Guo, Y., & Ye, J. (2019). Object detection in 20 years: A survey. Proceedings of the IEEE, 107(8), 1551-1577.