Deep Learning‑Based Approach for Automated Biometric Identification of Mugger Crocodiles

RCNN - Daksh Shah, Shalvi Modi, Sloka Thakkar

***Abstract*—This paper presents a machine learning–based framework for the automated biometric identification of free-ranging mugger crocodiles using unmanned aerial vehicle (UAV) imagery. The approach employs a pretrained model to generate bounding boxes that localize individual crocodiles in high-resolution frames. Subsequent feature extraction is performed using classical methods—including Scale-Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and Gray-Level Co-occurrence Matrix (GLCM)—to capture unique dorsal scute patterns. The experimental dataset consists of approximately 88,000 frames capturing 143 crocodiles across 19 locations in Gujarat, India. Preliminary results indicate that the proposed pipeline can serve as a non-invasive alternative for wildlife biometric identification. Future work will focus on ensemble learning strategies and the integration of multi-seasonal data to enhance the True Positive Rate (TPR).**

Key Words—Biometric Identification, Mugger Crocodiles, Machine Learning, UAV Imagery, Feature Extraction, Ensemble Learning.

I. Introduction

Biometric identification plays a vital role in wildlife monitoring and conservation. Traditional methods such as tagging can be invasive and stressful for animals. With the advent of UAV technology and advanced machine learning techniques, non-invasive biometric identification has become increasingly feasible. This work addresses the challenge of identifying individual mugger crocodiles in free-ranging environments using a classical machine learning approach. The motivation is to reduce human intervention and improve monitoring accuracy without the high computational costs often associated with deep learning models.

II. Literature review

Recent studies have demonstrated the benefits of combining deep learning and machine learning for biometric applications. For instance, Kokal et al. [1] reviewed integrated deep learning frameworks for biometric mobile authentication, while Khan and Bhatt [2] explored hybrid recognition systems using stacked autoencoders and Random Forest classifiers. More specifically, Ghosal et al. [3] applied a UAV-based CNN model (YOLO-v5l) to identify mugger crocodiles with an accuracy of 89.2%. Additional work [4]–[6] has extended these methodologies to other wildlife species, highlighting the potential of aerial imagery in automated species recognition. However, there remains a gap in exploring classical feature extraction and machine learning methods for such applications.

III. Dataset description

The dataset consists of 88,000 UAV-captured frames obtained using a DJI Mavic 2 Zoom drone. The captured dataset includes 143 free-ranging mugger crocodiles (approx. 1.5 m in length). It is collected from across 19 distinct sites in Gujarat, India.

Imaging Specifications:

- Frame resolution: 3840 × 2160 pixels at 96 DPI
- Optical zoom: 24–48 mm
- Flight height: 8–10 meters
- Recording duration: 30 seconds to 1 minute per session

Capture Protocol: When a crocodile was spotted ,the UAV was moved closer in order to obtain high-resolution images of the dorsal scute patterns. The OpenCV-Python package was used to process the captured video footage and extract individual frames for analysis.

IV. Methodology

In this work, we present a deep learning-based, robust pipeline for the detection and categorization of mugger crocodiles in drone-captured photos. Figure [X] shows the several sequential stages that make up the approach. The following is a description of our approach's main steps:

- YOLO-Based Object Detection: A pre-trained YOLO model is used to process an image that was taken by a drone at the start of the pipeline. This model creates bounding boxes around items it detects in order to locate and identify possible crocodile instances. Only the pertinent region of interest (ROI) is extracted for additional processing thanks to the bounding box generating stage.
- ResNet18 Feature Extraction: A pre-trained ResNet18 model is then used to extract features from the extracted ROIs. In particular, we make use of the last convolutional layer's output, which records high-level feature representations of the crocodiles that were found. Robust classification requires the acquisition of discriminative characteristics, which this step guarantees.

- FeatureNet for refinement:After the features have been retrieved, they are passed into the FeatureNet module, which processes and refines the feature representations to increase classification accuracy. FeatureNet analyzes the retrieved feature vectors and produces classification results by utilizing a trained deep learning model.
- Test Image Evaluation: The same object detection and feature extraction pipeline is used to input images during the testing phase. The trained FeatureNet model receives the extracted features and uses them to predict the detected object's class.
- Final Classification: The test image is finally classified using the FeatureNet output. The observed crocodile is reliably predicted by the trained model, which assigns the most likely class label.

This methodology ensures an efficient and accurate pipeline for detecting and classifying mugger crocodiles from drone imagery. By integrating object detection, deep feature extraction, and classification into a unified framework, the proposed system aims to enhance automated wildlife monitoring and conservation efforts.

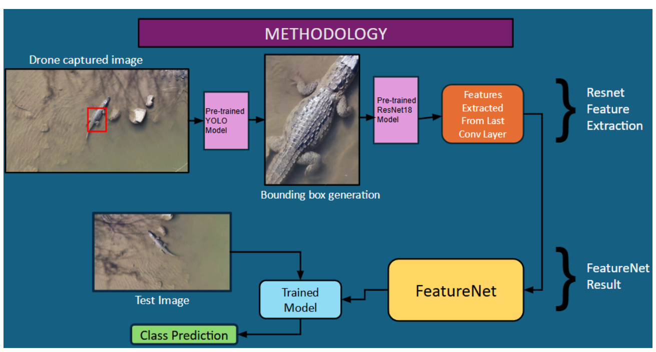

Fig 1. Flowchart of Approach

V. Experimental Results

Preliminary evaluations indicate that integrating ResNet-18 with FeatureNet effectively predicts data for a single season. Initial testing on new-season data will determine the model's generalization ability. While detailed quantitative metrics are still being refined, early observations suggest that if the model struggles with new-season data, a model merging approach—combining YOLO, ResNet, and FeatureNet into a unified framework—may enhance performance across varying environmental conditions.

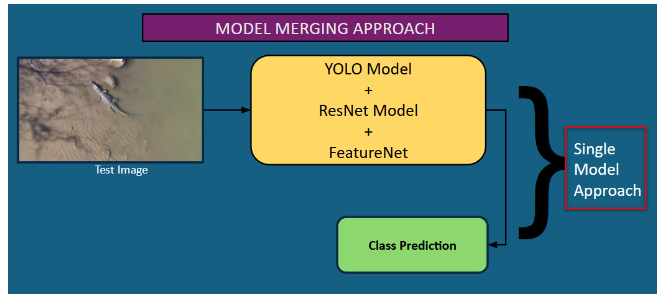

Fig 2. Model Merging Approach

VI. Results

Train Confusion Matrix

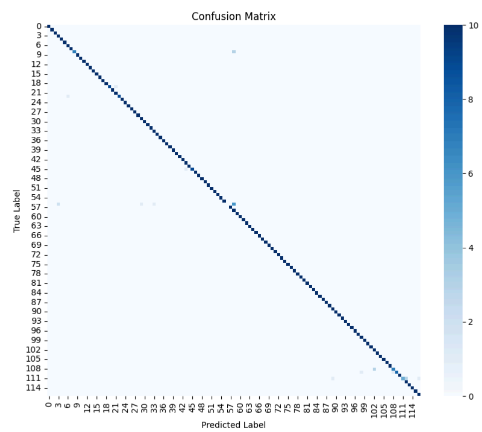

Fig. 3 Train confusion matrix

Figure 3 shows a strong diagonal pattern in the training confusion matrix, which indicates high classification accuracy for all labels. The majority of the projected labels match the actual labels, with very little misclassification. as the values in diagonal approaches to 10, it shows the highest frequency of accurate predictions across diagonals. Although they indicate rare misclassifications, sparse off-diagonal values have minimal impact. This suggests that the model has successfully picked up on the characteristics and trends found in the training set.

Test Confusion Matrix

The test confusion matrix (in Figure 4) maintains a similar diagonal dominance, confirming strong generalization performance. The highest frequency of correct classifications in the test dataset is lower (maximum of 4) compared to the training dataset. However, the minimal presence of off-diagonal elements indicates there are hardly any misclassifications.

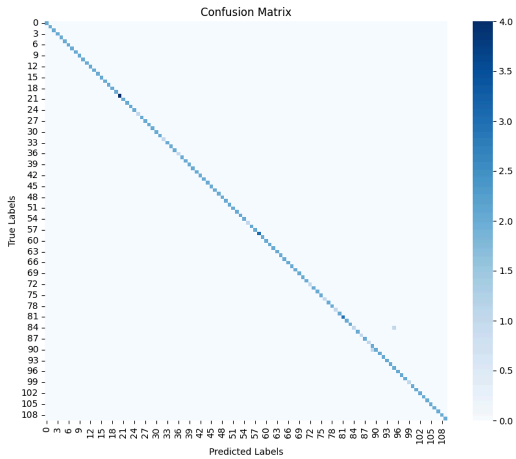

Fig 4. Test confusion matrix

Discussion

- Prediction confidence is higher in the training confusion matrix, as would be predicted given exposure to the training set.
- The model's capacity for generalization is demonstrated by the test confusion matrix, which shows somewhat lower confidence yet continuing to show a similar trend.
- Both matrices' sparse misclassifications imply that the model successfully separates labels in some classes with just slight confusion.

The model's strong classification accuracy and encouraging generalization to test data that hasn't been seen yet are confirmed by the confusion matrices overall. The remaining categorization mistakes can be reduced with more fine-tuning.

Conclusion

The confusion matrix analysis confirms that our model demonstrates strong classification performance with minimal misclassifications. The training results indicate a high level of learning, while the test results validate its generalization capability. Moving forward, we plan to implement a deep learning approach for further optimization. Our project will focus on two key aspects: (1) benchmarking results on the old single-season dataset and (2) improving the True Positive Rate (TPR) by incorporating multi-season data to enhance identification performance across various environmental conditions. Additionally, we will explore a model merging approach for future predictions to further refine classification accuracy.

References

[1] S. Kokal, M. Vanamala, and R. Dave, “Deep Learning and Machine Learning, Better Together Than Apart: A Review on Biometrics Mobile Authentication,” Journal of Cybersecurity and Privacy, vol. 3, no. 2, pp. 227–258, 2023, doi: 10.3390/jcp3020013.

[2] A. Khan and A. Bhatt, “Hybrid Biometric Recognition using Stacked Auto Encoder with Random Forest Classifier,” aging, vol. 4, p. 5.

[3] R. Ghosal, S. Shah, A. Patel, V. Patel, M. Raval, and B. Desai, “Identification of free-ranging mugger crocodiles by applying deep learning methods on UAV imagery,” Dryad, 2022, doi: 10.5061/dryad.s4mw6m98n.

[4] A. Delplanque, S. Foucher, J. Théau, E. Bussière, C. Vermeulen, and P. Lejeune, “From crowd to herd counting: How to precisely detect and count African mammals using aerial imagery and deep learning?,” ISPRS Journal of Photogrammetry and Remote Sensing.

[5] M. Moreni, J. Theau, and S. Foucher, “Do you get what you see? Insights of using mAP to select architectures of pretrained neural networks for automated aerial animal detection,” PLOS ONE, doi: 10.1371/journal.pone.0284449.

[6] A. Delplanque, “Multispecies detection and identification of African mammals in aerial imagery using convolutional neural networks,” Remote Sensing in Ecology and Conservation, Wiley Online Library, 2022, doi: 10.1002/rse2.234.