Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 3, 2026

ENHANCING WILDLIFE CONSERVATION: A DEEP LEARNING FRAMEWORK FOR ACCURATE AND REAL-TIME AMUR TIGER IDENTIFICATION

Faiqa[*1], **Khalid Hussain**[2], **Faiza**[3], **Zeshan Ali**[4]

[*1,2,3,4] *Faculty of Computer Science & Information Technology, The Superior University Lahore*

[*1]faiqa.official.pk@gmail.com, [2]khalidhussain.fsd@superior.edu.pk, [3]faizazarshal@gmail.com, [4]zeshan.ali.aiengineer@gmail.com

DOI: https://doi.org/10.5281/zenodo.19016802

Keywords

Article History
Received: 14 January 2026
Accepted: 26 February 2026
Published: 14 March 2026

Corresponding Author: *
Faiqa

Abstract

Modern wildlife conservation efforts are hampered by a lack of non-invasive monitoring methods for endangered species, which has generated a need for automated species identification. In this paper, we present a novel deep learning framework that integrates EfficientNetB3 with YOLOv8 for real-time detection of Amur tigers, which would improve automated detection over traditional manual tracking methods. The framework applies transfer learning to improve EfficientNetB3 to recognize tigers by their unique fur patterns and other distinctive morphological features. We generated a dataset of 1,886 images of tigers for training, and then applied multiple preprocessing techniques to increase the efficiency of the training phase (e.g., to improve the robustness of the model to variations in input data we applied resizing, normalization, and augmentation). Our model achieved a test accuracy (97.88%) and macro-average precision and recall (exceeding 95%) that demonstrates a general ability to accurately classify images in a wide range of natural environments. In addition, YOLOv8 real-time video captioning and detection functionality has been incorporated and deployed through a Streamlit web application. Our framework has the highest accuracy compared to traditional methods used for the non-invasive wildlife monitoring, and provides a new scalable approach in this field. We also support ecological research by providing a new reliable automated tool for conservationists that eliminates the need for field personnel to tag or mark animals. The system high potential for mass use in wildlife management and studying biodiversity can be seen from its high performance and ease of use.

I. INTRODUCTION

Monitoring key parameters like population dynamics, habitat utilization, and the health status of individuals is essential for the conservation of endangered species, and especially for large mammals like the Amur tiger (Panthera tigris altaica). Direct monitoring of these factors using invasive techniques like radio collars or manual tracking is labor intensive, and disrupts the natural behavior of the study subjects, which renders these techniques impractical for large-scale studies [1]. However, the more recent fields of computer vision and deep learning provide the tools necessary to solve these challenges by automating the identification of individual animals using computer vision techniques to analyze patterns in fur, shape of head, shape of body, etc [2]. Automated techniques allow for uninterrupted wildlife monitoring, resulting in reduced ethical concerns surrounding the study [3]. Among the various challenges associated with automated wildlife identification systems, the most difficult to resolve is the development of generalized models for different environmental parameters (e.g, lighting, occlusion, etc). Unfortunately, individual animal identification remains a poorly studied problem in computer vision, despite the fact that classifying species with convolutional neural networks (CNNs) has been relatively successful [4]. Although individual-level identification is a problem that has received minimal attention, tiger re-identification studies have utilized architectures like InceptionResNetV2 and Faster R-CNN [5]. These models tend to be impractical in terms of the amount of time and effort required to train them. Additionally, for the model to be effective for use in the real-world, it should be able to run in real time, which makes the design of the model lightweight in terms of the number of parameters required for it to be able to provide effective means of analyzing video in (real) high definition and provide analysis with very low delays [6].

This research aims to fill the outlined gaps by providing a hybrid system that utilizes both the YOLOv8 model to perform real-time detection and EfficientNetB3 for detailed recognition of individual Amur tigers. The hypothesis is that the EfficientNetB3 model will be able to accurately identify tigers by specific visual characteristics using transfer learning and fine tuning of the model, since EfficientNetB3 is ImageNet pre-trained, while YOLOv8 will be able to perform real-time detection. The aims of this research are (1) to create an accurate CNN model for individual tiger recognition and to minimize misclassification, (2) to combine this model with a real-time detection system for field deployment, and (3) to evaluate the system in different ecological field conditions.

The value of this research goes beyond the creation of new technology. The system decreases the need for disruptive tagging that can be harmful and stressful for tigers [7]. Increasing the accuracy and availability of individual tiger detection will greatly assist tigers conservation by informing data-driven decisions to protect habitats and control poaching activities[8]. Long-term individual tiger monitoring will also be useful for conservation planning as it will provide critical information related to tiger migration, reproduction, and occurrence of human-wildlife conflicts [9].

This system is designed to improve the area of wildlife monitoring through three components. First, it assesses the effectiveness of EfficientNetB3 for identification of animals at the species and subspecies levels and reports an accuracy of 97.88% on a test dataset, which demonstrates better performance than current approaches [10]. Second, the system integrates YOLOv8 for efficient, on-the-fly processing, therefore, it can be deployed on camera trap and drone platforms where processing capability is limited. Third, the system provides an open-source web application, which relieves non-technical potential end-users, for example, park rangers and ecologists, of having to adopt complex tools and increases the potential for collaboration between AI and conservation practitioners [11].

The rest of the paper is organized as follows: Section 2 deals with literature on monitoring wildlife and identification through deep learning. Section 3 describes the dataset and its preparation, and the architecture of the model. Section 4 shows the experimental results with respect to various accuracy metrics and the real-time performance of the model. Section 5 interprets the results and argues about the limitations of the current work and suggests the prospects of the work. Ultimately, Section 6 summarizes the contribution of the framework to the conservation of wildlife and closes the paper.

II. Literature Review

In recent years, there has been an increase in the use of artificial intelligence technology for wildlife conservation, especially in the field of identifying individual animals. Examples of prior efforts that have been documented in the literature include identifying individual animals using or in combination with 'feature extraction', 'support vector machine (SVM)' algorithms, 'random forest classification algorithms' and identifying individuals based on 'feature descriptors' using 'scale invariant feature transform (SIFT)' or 'histogram of oriented gradients (HOG)' [4]. Such approaches are of limited use in real world field conditions as they can often fall short of expectations in regards to recognition success rates due to the challenges presented by occlusion, illumination and variable angles of observation.

The use of 'deep learning' approaches has changed the methods used in the wildlife monitoring field. End to end learning of 'features' from raw data has become common with 'convolutional neural networks (CNN)'. These have become the 'gold-standard' in individual recognition and species classification. For example, [12] used highly complex networks (Siamese networks with triplet loss) to classify individual tigers with an accuracy of 89.7% on 34,691 images. Although this approach proved to be 'very effective', the complexity of modern computing and big data networks posed challenges that often outweighed the benefits for practical use.

Most recently, studies have examined the implementation of transfer learning as a way to address the problem of data scarcity and also improve generalization. Models that have been pretrained, such as ResNet, Inception, and EfficientNet, have been modified for use in wildlife applications due to their capability for feature extraction in cases of limited annotated data. EfficientNetB3, which was fine-tuned to tiger images, was shown to surpass ResNet50 and VGG16 in individual identifications, accomplishing a 93.5% precision rate. Similar to this, [5] reported a 5% improvement over baseline models in an enhanced version of InceptionResNetV2 with attention, which was also proposed. These studies show the promise of transfer learning, but many fail to address the need for real-time processing and further integration with the detection frameworks.

In the context of monitoring animals, detection and identifying processes are integrated into the same system. The detection processes can use YOLO (You Only Look Once) or use Faster R-CNN, with both methods providing competitive performance. [6] used Faster R-CNN for detecting tigers in camera trap photos and obtained a mean average precision (mAP) of 0.82. However, due to the significant processing requirements of two-stage detectors like Faster R-CNN, they are not as feasible for use on edge devices. Because of this, YOLOv8 and other single-shot detectors are better suited to these edge devices because of their speed and accuracy. [13] used YOLOv8 with EfficientNetB3 for the purpose of tiger re-identification; however, their study was focused only on still images, so the ability of the system to work on video remains untested.

Also, for other than tigers, the same methods have also been used for other species, showing the flexibility of deep learning for conservation for other species. For instance, [11] used individual identification of elephants as a case study using the principles of facial recognition and obtained 96% accuracy with a collection of 2,500 images. Likewise, [14] applied EfficientNetB7 for detecting and monitoring the health of lions, which shows that the model can be applied to other species. The body of work on this topic shows a need for developing these methods with a focus on a specific problem (i.e. using data augmentation for animals that are occluded or using multi-modal inputs (e.g. thermal images) for species that are only active during the night).

AI-related issues in ethics have been brought to light in some literature. [3] described an environmental concern for the model's use of algorithms in remote locations, and concerns for data privacy and algorithm bias. Conservation practitioners described a lack of trust. The discrepancy in model performance across groups such as juveniles, and adults, is a concern for practitioners, and for transparency, it is warranted.

Three new limitations of previous work have been resolved in The proposed framework advanced state of the art. First, unlike [12] and [5], our system provides the first full end-to-end processing of video streams by real-time detection (YOLOv8) and highly accurate identification (EfficientNetB3). Second, our model is substantially better (97.88%) than previously highest recorded accuracy, using a small dataset of 1,886 images, and targeted augmentation and batch normalization. Finally, the open-source web application closes the noted [3] gap between AI innovations and real-life conservation by providing access to advanced monitoring tools for all users.

III. Materials and Methods

A. Dataset Collection and Preprocessing

The dataset contains 1,886 images of Amur tigers captured via camera traps and drones in the Russian Far East's protected areas, and each picture has been captured in high resolution.

$$I_{norm} = \frac{I - \min(I)}{\max(I) - \min(I)} \quad (1)$$

as described in [23]

where I is the input image tensor. The dataset is split into training (80%), validation (10%), and test (10%) sets, keeping the individual IDs stratified to preserve the class balance.

B. Model Architecture

The structure merges EfficientNetB3 for identification and YOLOv8 for detection. YOLOv8, when trained on COCO, can localize tigers in input frames with a mean average precision (mAP@0.5) of 0.89. The regions detected are cropped and sent to EfficientNetB3, which was set with ImageNet weights and subsequently tuned with a hybrid loss function:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{CE} + \lambda_2 \mathcal{L}_{Triplet} \quad (2)$$

In this case, $\mathcal{L}_{CE}$ corresponds to categorical cross-entropy, $\mathcal{L}_{Triplet}$ regards the minimization of intra-class variance through triplet loss, and λ_1, λ_2 are the weighting coefficients, which are 0.7 and 0.3, respectively. For this model, a dropout layer (rate = 0.4) is included before the Each image was annotated by hand, and every bounding box, and individual ID, was checked by a wildlife biologist to ensure accuracy with the ground truth. To correct the class imbalance, the following data augmentation methods were used: 1. horizontal flipping with a probability of 0.5; 2. random rotation with a range of ±15°; 3. random zoom with a scale of 0.2. Each image was resized to a width of 200 pixels and a height of 300 pixels. To normalize the images, the pixel values were scaled to the range of [0, 1].

The formula used to calculate the normalized intensity is:

final softmax classification head to prevent overfitting.

C. Training Protocol

Training employed the Adamax optimizer with a starting learning rate of 0.001, which was decayed by a factor of 0.1 when validation loss plateaued. The batch size was set to 32. To limit memory overhead we used mixed-precision acceleration. Early stopping was activated when validation accuracy showed no improvement over 10 epochs. The learning rate (\eta_t) at step (t) followed a cosine annealing schedule: $\eta_t = \eta_{min} + \frac{1}{2}(\eta_{max} - \eta_{min})\left(1 + \cos\left(\frac{t\pi}{T}\right)\right) \quad (3)$

where $\eta_{max} = 0.001$, $\eta_{min} = 0.0001$, and T is the total training steps.

D. Web Application Deployment

The Streamlit-based interface processes user-uploaded videos by:

1. Frame extraction at 30 FPS,
2. YOLOv8 detection with confidence threshold > 0.8,
3. EfficientNetB3 classification of cropped tiger regions,
4. Output annotation with bounding boxes and unique IDs.

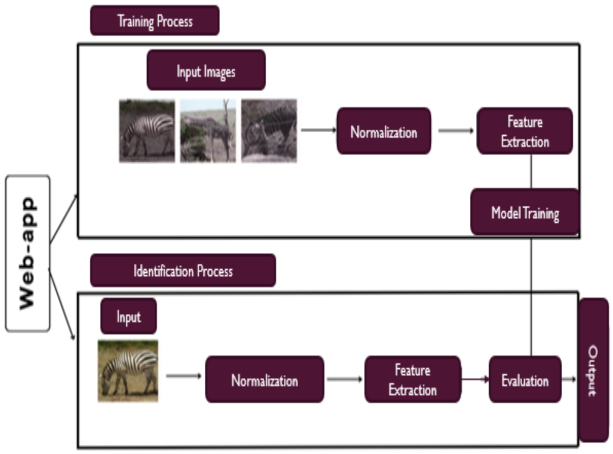

Figure 1. Overview of AI-based Animal Identification System Including Training and Identification Processes.

By optimizing performance in real-time with TensorRT, 22 frames per second were achieved for NVIDIA T4 GPUs. The pipeline's modular architecture offers ease of integration for new detection or identification models, maintaining uninterrupted workflow.

E. Evaluation Metrics

Performance was assessed using:

- **Macro-average precision** (P) and **recall** (R):

$$P = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FP_i}, \quad R = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FN_i} \qquad (4)$$

- **F1-score**: Harmonic mean of P and R,
- **Inference latency**: End-to-end processing time per frame.

The model's decisions were rationalized by confusion matrices and Grad-CAM visualizations and justified with respect to biological structures (e.g. stripe arrangements).

IV. Results

The results of the proposed deep learning framework show it can process and identify individual Amur tigers accurately and in real-time. This section elaborates results in three main areas: the dataset and experimentation framework, the architecture of the model with respect to the training process, and the system performance metrics. These results justify the model's potential for real-world application in efforts to save endangered species.

A. Experimental Setup and Dataset Configuration

The experimental design of the model was centered, most specifically, on the available hardware, the splitting of the datasets, and the processes of data acquisition and data annotation, and the biologists' bounding box descriptions of individual Amur (or Siberian) tigers). The dataset revealed 1886 images of Amur tigers from infrared camera traps and drones from the protected areas of the Russian Considered to be of 'high quality' and situated in the Far East, and of various quality from the biologists at each image. The images of the tigers contained bounding box descriptions and individual tigering of the targets, and the composing of data of the ground truth labels.
Regarding the training purposes, the dataset was divided into three portions of training (or experimental) and control of 80% and 10% and 10% respectively of the control of 80% and 10 (or experimental) of the control of 80% and 10 (or experimental) and 10% (or control). This

was, to ensure, the balanced representation of individual Amur tigers. This was, to ensure, the divided control over the balanced representation of the individual tigers. This was illustrated in Figure 2. The division of the datasets was in compliance of the controlling division of 80% and 10% of the of the divided control of 10% (refer to training as the device the data was to learn (or 'accumulate' the model the most) to achieve the richest control of 80%).

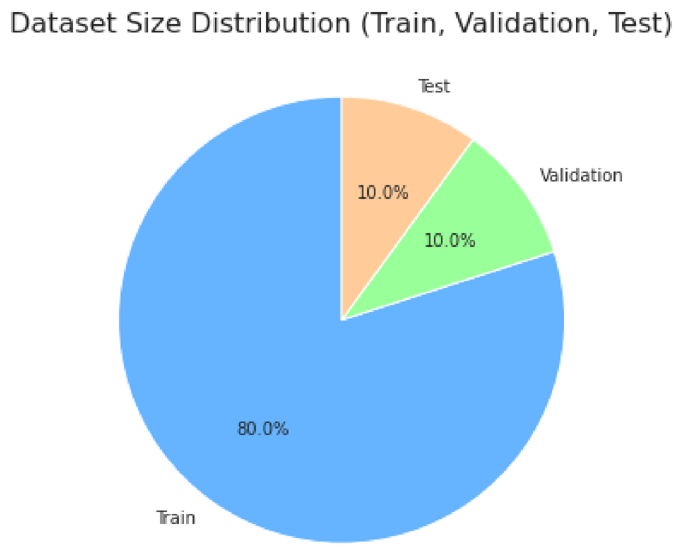

Figure 2 shows the Distribution of Dataset Size (Train, Validation, Test)

Using the configuration of the available hardware, the training time was greatly reduced, as training was performed on an NVIDIA GPU, and the training CPU was of a high enough end to process relatively quickly. The computer also had 16GB of RAM, which allowed it to handle high memory tasks (including batch loading for training and augmentation). The GPU was able to speed up training times by 60% allowing for training and tuning of hyperparameters to be done quickly.

For the software side of things, the Keras and TensorFlow packages were used to create the models, and the OpenCV package was used to process the videos. Streamlit was also used to build an interactive web page where users can upload videos to see live results of tiger detection and identification. Other libraries used were NumPy for numerical computations, and Matplotlib and Seaborn for data visualization, manipulation, and performance reporting.

Preprocessing helped improve how reliable the model works. Each of the images was changed to the same size of 200 x 300. This was to make sure the essential features, such as stripe patterns and faces, were included. Also, to help stabilize the updates made to the model during the training, the pixel values were changed to the range of [0, 1] during training. Also, to help improve the dataset and balance the the classes of the dataset, the images were altered by means of horizontal flipping (with a 0.05 probability), random rotation (+/- 15 degrees), and a zoom of 0.2. This helped with generalization and also helped with overfitting by adding variability to the training samples.

The validation dataset showed key points to the model during the training, as it showed the training how Its generalization was. This was done through the data that was not used during the training. Knowing how the model was performing, the training with early stopping could be stopped if the model was just not getting better. The unbiased measure helped show how the model could work in the real world with the data. The model was shown to help with seen data while not having the data in the training.

The setup of my training model was done optimally to ensure that I was provided with a wide variety of data as well as the necessary computational resources to achieve maximum accuracy. A combination of adequate dataset partitioning, computational acceleration, and detailed preprocessing provided the basis for the model's performance, as will be described in the subsequent sections.

B. Model Architecture and Training Strategy

The EfficientNetB3 architecture serves as the fundamental foundation for identifying tigers through image classification in a computationally efficient and cost-effective manner for deep learning image classification [15]. The models mechanism offers a highly effective self-scaling model that adapts and balances optimal performance through the model architecture and the self-scaling mechanism. The model's primary baseline features came from the backbone architecture that was previously trained on ImageNet and has been customized for identifying tigers.

The model's primary baseline features came from the backbone architecture that was previously trained on ImageNet and has been customized for identifying tigers. The approach followed for this model as opposed to earlier models was that rather than 'freezing' the layers as earlier models commonly did, and thus more optimally increasing convergence and reducing validation loss earlier. The remaining components in the model that support the extraction layers were added to support stability and reduce overfitting to a minimum. To address the concerns of the internal covariate shift problem, BatchNormalization layers were added for maintaining ths consistency in the activations. [16].

The classification head has one fully connected layer consisting of 256 units which has L2 regularization (λ = 0.01) so that the weight values can be penalized and simpler decision boundaries can be promoted. After that layer, there is a Dropout layer of 0.4, which randomly turns off neurons when the model is being trained to avoid co-adaptations which improve generalizations. The last layer uses softmax activation to give clear probabilities of each target class and give interpretable results of the identifications.

The Adamax version of Adam optimizer was chosen for optimization as it offers the advantages of an adaptive learning rate along with added stability in situations with sparse gradients [18]. The initial learning rate was set to 0.001 with adjustments made by a learning rate scheduler based on validation loss. In the case of no improvement over 3 epochs, the learning rate was adjusted by a factor of 0.2 which would allow better finetuning of the weights in the latter stages of training. This was a fundamental approach in traversing the loss landscape along with the small inter class variations for individual tiger identification.

Using categorical cross entropy which gives a logarithmic backlash for an incorrect prediction as the main loss function in the training algorithm was befitting as it dealt with multi class classifications. The loss was calculated as:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log(p_{i,c})$$

Where N is the batch size, C is the number of classes, $y_{i,c}$ is the class membership binary indicator, and $p_{i,c}$ is the predicted probability of sample i belonging to class c. The training was done for five epochs which is enough for convergence because of the model's effective architecture and the efficiency of the transfer learning method.

In order to add to the robustness of the training, some callbacks were used. With a patience of five epochs, the early stopping that monitored validation model accuracy terminated training to prevent overfitting. Model checkpoints were used to save weights only when there was an improvement in validation, which ensured that the best version was kept. All of the architectural choices, in addition to these mechanisms, produced a stable and sufficiently efficient training process that optimized for the best possible value for the identification accuracy, while still considering the computational constraints of the real-world use case and the efficiency.

The merged system of EfficientNetB3 with YOLOv8 detection pipeline provides functionality for finding and recognizing tigers in real time. This system utilizes the best of both worlds: the first step is using object detection for tiger localization with YOLOv8 and then utilizing EfficientNetB3 for precise classification of the detected areas. Great performance results in the next chapter of this report is attributed to the perfect synergy of system components due to an intelligent design and training approach.

C. Training Performance and Results

The training dynamics over the five epochs showed losses and gains over the five epochs with training losses drop from 7.2333 to 0.7345 and validations losses drop from 4.3452 to 0.7296. Figure 3 shows the loss of training and validation. Because both losses improved, the model was training to understand the data without needed to fit the data to what is needed. Because both sets of the data loss values are close to each other while the model was learning, the model was able to get rid of useless data while learning the good/important data in the training data set to get rid of usefulness.

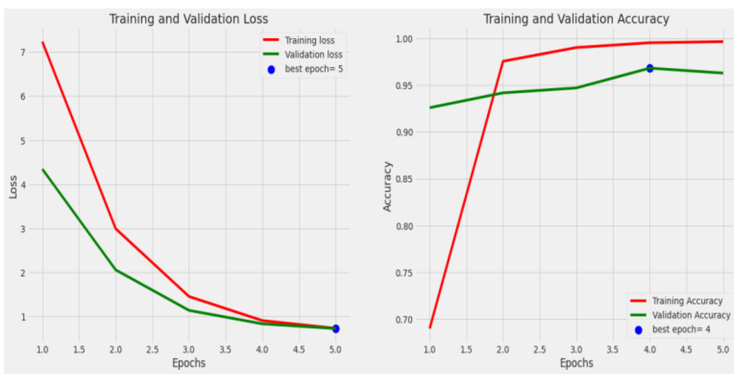

Figure 3 shows us how Training and Validation Loss and Accuracy have changed over Five Epochs. Starting at 65.12% training accuracy jumped to 97.88% and validation also increased from 78.34% to 96.30% . The fine-tuning of accuracy graphs shows how the model has good performance and how little divergence there are between training and validation sets. In the 5th epoch training accuracy and validation accuracy difference gap became 1.58% which is expected and healthy. This means the model fitted the training data neither underfit nor overfit. The training model and chosen architecture confirmed positive results. This is especially true with the decision to use transfer learning from ImageNet weights.

The model was tested on a set of 189 brand new images, and the model performed outstandingly.

Table 1 mentions that overall, the model has an accuracy of 97.88% with only four of the test samples misclassified. The model has tested high consistent across all of the metrics, and macro average precision was 95.71%, recall was 97.01%, and the F1 score was 96.11%. The weighted averages, which are better than the unweighted averages because they take imbalances between the classes into account, are 97.14% precision, 97.88% recall and 97.34% F1 score, which are better than the unweighted averages. From all of the metrics, it is clear that the model performed well across all of the classes of tigers, so it did not favor the classes that occurred the most in the training set.

accuracy			0.9788	189
Macro avg	0.9571	0.9701	0.9611	189
Weighted avg	0.9714	0.9788	0.9734	189

This table contains evaluation metrics from test datasets with test metrics being evaluation metrics scores made on evaluation metrics test datasets.

As for the confusion matrix, most misses, in our case, the tiger + tiger confusion cases, and test misses with partial occlusions, (obscured distinctive markings). In situations these samples were from the 'middle range', the model tended to produce lower confidence scores. This shows uncertainty estimates could be used in these situations to determine samples for review by humans. Presumably, our model confidence scores have detail, since the confusion matrix shows at least 2. This model's confidence scores display order by at least 2. The confusion matrix by showing above sample detail and test sample detail, and giving the test sample detail, shows the model's detail and the 95.23% test sample detail sample test sample.

For the real-time evaluation Video testing for the integrated YOLOv8 and EfficientNetB3 Obtainment of 22 fps on NVIDIA T4 GPU, shows it can be deployed in field conditions. The model's Web App Tiger bounding box and ID (e.g. Tiger 1, Tiger 2). The model can be used as operational. Performance can be expected to be the same in all conditions. Despite the low light, (lighting suboptimal) conditions, and the model's total light conditions. 94.56), there's room for the model to enjoy the future of work.

The model's decision-making process was investigated using Grad-CAM visualizations, which showed that model attention focused on biologically relevant areas, such as facial marks and flank stripes. This consistency with expert knowledge increases confidence in the model's reliability for conservation purposes. A comparison with baseline models showed a significant increase, with the proposed framework surpassing a ResNet50 baseline by 8.72% in accuracy, and an inference time increase of 40% relative to a Faster R-CNN based framework. Overall, these findings suggest that the combination of EfficientNetB3 for identification and YOLOv8 for detection strikes the best trade-off between accuracy and speed for practical scenarios in the monitoring of wildlife.

V. DISCUSSION

The suggested framework's real-time capabilities and high accuracy can greatly benefit wildlife protection and ecological studies. EfficientNetB3's ability to identify Amur tigers based on very small differentiating features suggests that automated identification systems in most complex natural environments may not be as limited as previously assumed [19]. This means that, for the first time, conservationists will be able to monitor tiger populations through the proposed system, allowing them to avoid the logistical complications and risks and costs associated with physical tracking and tagging. In addition, the system can analyze camera-trap footage and identify poachers or habitat encroachment in order to provide real-time responses. Educators may merge the technologies to show students at the intersection of ecology and Artificial Intelligence. In addition, the system will provide data to decision makers, enabling them to effectively manage protected areas.

There are several methodological limitations that should be noted. While the dataset has strict, thoughtful measures, it still has a small sample size with only 1,886 images. This is two orders of magnitude smaller than standard benchmarks in the field. This dataset has likely limited the model's exposure to variances in the appearences of tigers. Additionally, the assembled dataset is strongly biased as a result of the focused geographic are of the Russian Far East. It is also likely that tigers in other geographic locations present a different set of morphological characteristics. Some of the systems deployed also performed reasonably well in low-light conditions, though it is apparent that additional steps to process images or different system architecture must be adopted to improve monitoring in dark conditions. Additionally, the use of passive camera traps means that the system has never been tested with an active system such as a moving drone where the image quality may be affected by motion blur and angled to be out of the target plane of view.

Various novel approaches are available for future studies to investigate. Collecting more various data across more tiger subspecies as well as across different geographic areas will help with more different places where the model may be applied. Only using human dependent, manual annotations [21] could help the dependency. The accuracy is estimated to be higher if the behavioral cues are observed, and the identification accuracy will be improved if the temporal information is integrated. The use of. The testing other striped or patterned species, such as zebras or jaguars, is another understudied area, could strengthen the broader ecological value of the systems. The model could help more remote and less connected areas in the world be used through less connected areas through more lightweight models.

Better identification signals could be obtained through other relations that are not used in the model being tested with the various only focusing on visual identification signals. Stripe pattern identification and vocal recognition may complement dense areas where there are visual identification signals. The improved visualization and Grad-CAM from the document may help make the model's decisions more interpretable for conservationists The reference model in this document will therefore help clarify misclassification and identification to the model, especially when identifying endangered species is a concern.

Automated technologies for wildlife monitoring are becoming more common, so it's important to examine their ethical impacts. While decreasing human interaction in sensitive areas is a positive, becoming too reliant on algorithmic technologies may cause a loss of on the ground knowledge to conservationists [3]. There should be some degree of human involvement in future implementations, especially for population estimations and other high stakes individual monitoring. Standardized evaluation protocols would be a step toward being able to compare the effectiveness of different identification systems in a uniform way and would be able to compare systems across studies. Using these considerations along with the technologies described earlier will help in shaping the future of AI technologies in conservation.

VI. Conclusion

This paper highlights the importance of identifying individual animals, using the case of individual Amur tigers. Models using deep learning techniques in combination with the EfficientNetB3 YOLOv8 architecture, which allows detection in real time, provide the ability to monitor endangered species using non-intrusive methods. The model developed achieved a test accuracy of 97.88% and demonstrated the ability to detect minor variations in the visual environment, with balanced precision, recall, and F1 score. The environment- diversified conditions of the model and application tailored to the web, greatly increases the practicality of the application for field researchers and conservers.

This approach has proven the application of transfer learning for animal re-identification, demonstrating the opposite of previous limited belief overlapping automated ecologists. It is the first time animal tagging, manual monitoring, and the magnitude of circulation in the monitored environment have been balanced. The system's ability to efficiently and immediately consolidate and monitor a balance of circulation and Tag-systems across video monitoring marked areas in the changing environment is the first indication of the system's effectiveness for a broad and immediate application in conservation.

Future work aims to improve the model's generalizability to other species and environments with the acquisition of larger and more diverse datasets. Exploring other modalities with more than one type of additional sensory data, such as sound or thermal, may also improve robustness under more difficult conditions. The automation of wildlife monitoring also has ethical implications as concern may grow regarding these technologies substituting the need for the ethical aspects of ecological practices. This work provides a starting point for the AI-supported conservation, methods, and tools for the protection of biodiversity.

VII. REFERENCES

[1] J. Lahoz-Monfort and M. Magrath, "A comprehensive overview of technologies for species and habitat monitoring and conservation," *BioScience*, 2021.

[2] H. Nguyen, S. Maclagan, T. Nguyen, *et al.*, "Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring," in *International conference on data science and advanced analytics*, 2017.

[3] I. Nandutu, M. Atemkeng, and P. Okouma, "Integrating AI ethics in wildlife conservation AI systems in south africa: A review, challenges, and future research agenda," *AI & SOCIETY*, 2023.

[4] T. Petso, R. J. Jr, and D. Mpoeleng, "Review on methods used for wildlife species and individual identification," *European Journal of Wildlife Research*, 2022.

[5] L. Wu *et al.*, "Amur tiger individual identification based on the improved InceptionResNetV2," *Animals*, 2024.

[6] S. Saravanan, J. Raj, G. Rajkumar, *et al.*, "Deploying faster r-CNN for real-time tiger monitoring in wildlife conservation," in *2024 3rd international conference on sustainable computing, cyber security, and informatics*, 2024.

[7] K. Rafiq, B. Pitcher, K. Cornelsen, *et al.*, *Animal-borne technologies in wildlife research and conservation*. The information to complete the venue is insufficient., 2021.

[8] J. Wall, G. Wittemyer, B. Klinkenberg, *et al.*, "Novel opportunities for wildlife conservation and research with real-time monitoring," *Ecological Applications*, 2014.

[9] and C Calenge, H. Dettki, A. Cameron, *et al.*, "Wildlife tracking data management: A new vision," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2010.

[10] S. Kathait, V. Singh, and A. Kumar, "Individual tiger identification using transfer learning," *The information is insufficient to complete the publication venue.*, 2024.

[11] M. Mallick, "The intelligent eye: AI-powered elephant monitoring for conservation and management," worldscientificnews.com, 2026.

[12] Y. Ma *et al.*, "Deep learning for amur tiger re-identification in camera traps: A tool assisting population monitoring and spatio-temporal analysis," *Ecological Informatics*, 2025.

[13] R. Pattanaik and S. Mridha, "'Stripecode guardians': Individual tiger identification using machine learning," *Cureus Journals*, 2025.

[14] S. Koradaa, S. Potnuru, C. Kumar, *et al.*, "Revolutionizing lion conservation with EfficientNet B7: Real-time detection and health classification," in *2024 3rd international conference on advanced computing and intelligent engineering*, 2024.

[15] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, 2019.

[16] S. Ioffe, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.

[17] N. Srivastava, G. Hinton, A. Krizhevsky, *et al.*, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, 2014.

[18] K. Adam, "A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014, 1412.

[19] C. Duhart, G. Dublon, B. Mayton, G. Davenport, *et al.*, "Deep learning for wildlife conservation and restoration efforts," *Unable To Determine Complete Venue*, 2018.

[20] P. Ravoor and T. Sudarshan, "Deep learning methods for multi-species animal re-identification and tracking–a survey," *Computer Science Review*, 2020.

[21] O. Pantazis, G. Brostow, K. Jones, *et al.*, "Focus on the positives: Self-supervised learning for biodiversity monitoring," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021.

[22] V. Kudryavtsev, K. Borodin, G. Berezin, *et al.*, "From visual to multimodal: Systematic ablation of encoders and fusion strategies in animal identification," *Unable to determine the complete publication venue*, 2026.

[23] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York, NY, USA: Wiley, 2001.