

# Revolutionizing Tiger Detection: A Deep Learning and Frame-by-Frame Analysis Approach

Karnam Nithin  
School of Computer Science Engineering  
Vellore Institute of Technology  
Chennai 600127, India  
karnam.nithin2021@vitstudent.ac.in

Gajjala Sai Charan Reddy  
School of Computer Science  
Engineering  
Vellore Institute of Technology  
Chennai 600127, India  
gajjalasai.charanreddy2021@vitstudent.ac.in

V Pattabhi Ramanna  
School of Computer Science  
Engineering  
Vellore Institute of Technology  
Chennai 600127, India  
vullipattabhi.ramanna2021@vitstudent.ac.in

G Karthik  
School of Computer Science  
Engineering  
Vellore Institute of Technology  
Chennai 600127, India  
@vit.ac.in

Dr. Sudheer Kumar E  
School of Computer Science  
Engineering  
Vellore Institute of Technology  
Chennai 600127, India  
sudheerkumar.e@vit.ac.in

**Abstract**—Tigers, majestic predators, face numerous threats such as poaching, habitat loss, and human-wildlife conflict. Effective monitoring is crucial for their conservation, but traditional methods like camera traps and field surveys are time-consuming, labor-intensive, and prone to human error. In this work, we propose a novel approach utilizing deep learning models, specifically the Inception and ResNet architectures, for automated tiger detection in photos and videos. By training these models on extensive datasets of tiger images, we enable them to analyze individual frames with exceptional accuracy and efficiency, allowing for real-time analysis of large video datasets. Our system offers significant advantages over traditional methods, reducing manual effort and human error, enabling real-time analysis of camera trap data for faster responses to potential threats, and providing scalability for deployment across vast areas. The data generated by our automated tiger detection system can inform conservation efforts, optimize resource allocation, and ultimately contribute to the protection of these magnificent creatures. In summary, our work presents a cutting-edge solution for tiger monitoring, harnessing the power of deep learning to overcome the limitations of traditional approaches and provide conservationists with a powerful tool to aid in their efforts to protect and preserve these iconic predators.

**Keywords**—Tiger conservation, Automated tiger detection, Camera traps, Inception architecture, ResNet architecture, Human-wildlife conflict, Conservation strategies, Population analysis, Wildlife protection

## I. INTRODUCTION

Tigers, as critical hunters, assume a significant part in keeping up with environment balance, adding to biodiversity and the solidness of their territories [1]. In any case, these great animals face a huge number of dangers, including poaching, environment misfortune, and human-untamed life struggle, requiring worldwide endeavors for their preservation [1, 7, 17]. The desperation of these endeavors is featured by the intricate difficulties tigers face, making their conservation a worldwide goal.

Traditional observing strategies, depending on work serious procedures, for example, camera traps and field studies, have been the foundation of natural life preservation [1,11]. Nonetheless, these strategies are not without their deficiencies. They are inclined to postponements and mistakes, preventing opportune mediations and viable protection methodologies [1,15].

Perceiving the restrictions of these customary methodologies, this work proposes an imaginative arrangement: utilizing profound learning models for computerized tiger location in photographs and recordings, with a particular accentuation on continuous observing. The squeezing need for powerful and adaptable arrangements emerges from the declining tiger populaces, basically ascribed to different dangers [11,28].

Manual ID strategies, for example, investigating camera trap pictures, spoor ID (creature tracks), and pugmark reviews (following paw prints), while significant, are hampered by innate limits. These incorporate onlooker predisposition, weakness, and the potential for mistakes, compromising the general precision of the gathered information [1,17].

To address these difficulties, a state-of-the-art framework has been created, using progressed profound learning models, explicitly Inception and ResNet [4,10,11]. These models are prepared on broad datasets, empowering them to perform exact examination at the singular casing level in video film [5,18].

The advantages of this proposed profound learning arrangement are complex. It altogether diminishes the manual exertion and human mistake related with customary strategies, prompting more exact and solid outcomes. The continuous investigation capacities of the framework empower quick reactions to potential dangers like poaching or human-natural life struggle, improving generally speaking preservation techniques [10,12]. Besides, the adaptability of

the framework considers organization across immense regions, giving important bits of knowledge into tiger populaces, dispersion examples, and living space utilization [11,14].

Drawing matches with the field of clinical examination, especially skin malignant growth forecast [6], features the unprecedented capability of computer based intelligence procedures. The utilization of profound learning models for tiger identification connotes a change in outlook in untamed life observing, reflecting the achievement found in skin malignant growth recognition. The combination of cutting edge models, like Convolutional Brain Organizations (CNN), guarantees uplifted accuracy and effectiveness in the distinguishing proof and following of tigers [10,11].

This paper denotes the initiation of an exhaustive survey expecting to evaluate the dependability and security of Computerized reasoning/AI (man-made intelligence/ML) innovations custom-made explicitly for mechanized tiger location [11,29]. While essential consideration settings stay urgent for early mediation in clinical findings, with regards to natural life observing, the center movements to different regions, especially those in danger because of anthropogenic exercises [17,31]. Resulting segments of this paper will carefully investigate the nature of accessible proof, the formative phase of man-made intelligence/ML advancements, distinguished proof holes, and the possible use of these state of the art advancements in the preservation of tigers, adjusting intimately with the organized structure.

In this pursuit, it is fundamental to look at and comprehend the progressions and difficulties presented by simulated intelligence/ML advancements in robotizing tiger location. The examination will give important experiences into the capability of these advancements to reform natural life preservation techniques and make significant commitments to the security of these eminent animals [32,33].

## II. LITERATURE REVIEW

According to Rolnick et al. (2017), a deep neural network can learn from many noisy labels. They investigated how set size and learning rate affected the performance of the model. Van Horn et al. (2015) argue that as long as the error rate is not too high, learning methods based on CNN features and part localization are robust to annotation errors and damaged training data.

Yuan et al. (2018) recognize that the diversity of these networks is crucial during this joint evaluation process, so that different networks can provide different insights or information about the input sample. This can improve the performance of joint estimation. There are two ways to overcome the challenge of learning from noisy labels. Assuming that clean labels are not available, the first method involves learning directly from noisy labels. Yuan et al. (2018) observed and improved deep neural network and learning using a set of 5000 clean samples.

Tabak et al. (2019) explores the use of machine learning to automatically identify animal species in camera trap images.

Manually analyzing these massive datasets is time-consuming. The authors trained a model on millions of images, achieving high accuracy (98%) for species in the training data and good performance (82%) on unseen data from Canada. They also developed an accessible software package for ecologists to utilize this technology in their studies. This approach has the potential to significantly improve efficiency and data collection in ecological research. To correct for noisy input labels, Veit et al. (2017) developed a tag cleaning network moderated by a number of clean tags. Han et al. (2018), Jiang et al. (2017) and Ren et al. (2018) all used pure validation sets for strength training data.

Greenberg et al. (2020) investigates limitations of deep learning for identifying species in camera trap images. While promising, these systems face challenges: limited training data, messy and imbalanced images, and performance decline in unseen locations. The authors test six deep learning models on a dataset with moderate image numbers (47,279) and various environments. They find DenseNet201 performs best with 95.6% accuracy in "trained" locations, highlighting potential for smaller scale ecological studies, but emphasizing the need for further research on generalizability.

Swanson et. Al describes the creation of a valuable resource for studying African savanna mammals. The authors deployed camera traps across a vast area of Tanzania, capturing millions of images of 40 mammal species. To analyze this massive dataset, they enlisted the help of citizen scientists through the website Snapshot Serengeti. Volunteers classified the animals in the images, allowing researchers to efficiently gather data on species presence, abundance, and behavior. This unique dataset offers valuable insights into African savanna ecology and demonstrates the power of citizen science in ecological research.

Vankdothu et. Al(2019) proposes a system for brain image recognition on the Internet of Medical Things (IoMT) using a combination of adaptive feature selection and an Entropy-based Deep Neural Network (EDNN). Early and accurate diagnosis of brain conditions is crucial, and this system aims to achieve this goal within the resource constraints of IoMT devices. The system works in four stages: pre-processing, feature extraction, adaptive feature selection, and finally, classification using the EDNN

Tuia et. Al (2022) highlights machine learning's growing potential for aiding wildlife conservation efforts. Data Collection - ML analyzes camera traps, remote sensing, and citizen science data for monitoring animal populations and habitats. Analyzing Complex Patterns helps uncover hidden ecological relationships and predict future trends . Scaling Up Efforts enabling analysis of vast datasets and real-time decision-making.

Khalajzadeh et. Al Used for feature extraction. CNNs are adept at automatically learning relevant features from images, making them ideal for tasks like face recognition where identifying distinguishing characteristics is crucial. Logistic Regression Classifier: Employed for classification. This simple yet effective method takes the features extracted by the CNN and classifies them to identify the corresponding

individual. leveraging CNNs for robust feature extraction and a simpler classifier for final identification, the system aims to achieve good recognition rates while maintaining computational efficiency.

This article by Okafor et al. (2016) compares two methods for recognizing wild animals in images: deep learning and bag-of-visual-words. They found that deep learning models achieved significantly higher accuracy than the traditional bag-of-visual-words approach, suggesting deep learning as a more promising technique for automated wild animal recognition.

Premarathna et. Al p.591 outlines a system that uses Convolutional Neural Networks (CNNs) to detect elephants in images. The system aims to reduce human-elephant conflict by automatically identifying elephants approaching human settlements and potentially triggering alert systems. This approach could help protect both humans and elephants by creating early warnings.

Zmudzinski et. Al (2018) explored using deep learning for guinea pig image classification. They employed Nvidia DIGITS and the GoogLeNet model, successfully distinguishing between three fur types: skinny, Abyssinian, and crested. Notably, including empty background images improved the model's accuracy. This research paves the way for utilizing deep learning in animal recognition for various applications.

Huang et. Al propose a deep learning platform for bird image retrieval and recognition. Their system utilizes a convolutional neural network (CNN) trained on a dataset of bird images from Taiwan. The CNN incorporates techniques like skip connections to improve feature extraction and achieves high accuracy (98%) in recognizing 27 bird species. This method offers a potential tool for birdwatchers and researchers to efficiently identify birds from images

Allken et al. (2019) present a method for fish species identification using a convolutional neural network (CNN) trained on synthetic data. This addresses the challenge of limited real-world fish image datasets. Their approach achieved 94% accuracy for specific fish species, demonstrating the potential of synthetic data for training CNNs in situations with real data scarcity.

Lai et al. (2019) present a dog identification system using a combination of "soft biometrics" (breed, height, gender) and neural networks. They found that incorporating this information with facial features through a decision network improved identification accuracy compared to using facial features alone (84.94% vs 78.09%). This approach may enhance dog identification in various applications, like lost pet recovery.

### III. METHODOLOGY

#### I. Data Collection

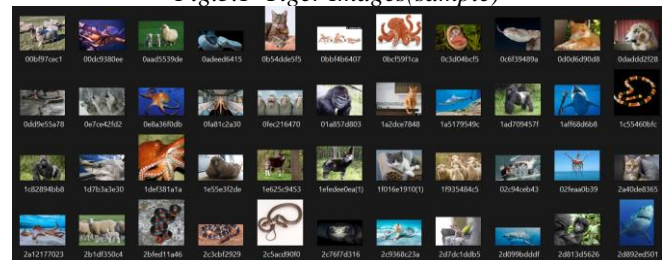
**Source Selection:** The foundation of any machine learning project lies in the quality and diversity of the dataset. In our tiger detection project, we meticulously selected data sources from various reputable outlets. Wildlife reserves provided authentic images captured in the tigers' natural habitats, offering valuable insights into their behaviors and interactions with the environment. Additionally, contributions from conservation organizations and research institutions enriched the dataset with scientifically collected data obtained through field surveys and monitoring efforts. The inclusion of images from online repositories such as Images.cv and academic databases further diversified the dataset, ensuring a comprehensive representation of tiger populations across different geographical regions.

**Image Acquisition:** Capturing images of tigers necessitated the utilization of multiple techniques, each offering unique advantages in data collection. Camera traps, strategically positioned in tiger habitats, autonomously captured images, providing researchers with a non-intrusive method to observe tigers in their natural settings. Field surveys, involving direct observation and photography during research expeditions, supplemented the dataset with ground-truth data collected by field experts. Additionally, remote sensing technologies like drones and satellites facilitated the acquisition of high-resolution aerial imagery, covering vast geographical areas and providing a broader perspective on tiger habitats and distributions.

**Annotation Process:** Annotating images is a critical step in dataset preparation, ensuring that machine learning models can accurately distinguish between tiger and non-tiger instances. Trained annotators meticulously labeled each image, marking the presence or absence of tigers based on identifiable features such as stripes, size, and morphology. To maintain data integrity and consistency, quality control measures such as inter-rater reliability checks and consensus-based labeling were implemented. These measures minimized labeling errors and discrepancies, ensuring the reliability of the annotated dataset.



*Fig.3.1 Tiger Images(sample)*



*Fig.3.2 Non-Tiger Images(sample)*

#### II. Data Preprocessing

**Image Resizing:** Standardizing the size of images to a uniform dimension is crucial for ensuring compatibility with deep learning models' input requirements. While ResNet-50 V2 typically requires images to be resized to 224x224 pixels, InceptionV3 utilizes a slightly larger input size of 299x299 pixels. Resizing images simplifies computational processes and streamlines model training by eliminating variations in image dimensions.

**Normalization:** Normalizing pixel intensity values to the range [0, 1] enhances model convergence during training. By scaling pixel values, normalization mitigates the impact of variations in pixel intensity distributions across images, ensuring consistent data representation and improving model performance.

**Data Augmentation (if applicable):** While not utilized in this project, data augmentation techniques such as rotation, flipping, and brightness adjustments can be applied to increase the diversity of the training dataset. Augmentation introduces variations in image orientation, lighting conditions, and perspectives, enabling the model to learn robust features that generalize well to unseen data.

### III. Model Selection

**Architectural Overview:** The choice of deep learning architectures plays a pivotal role in the success of our tiger detection task. ResNet-50 V2 and InceptionV3 were selected for their advanced architectures and proven performance in image classification tasks. While ResNet-50 V2 is known for its deep residual learning architecture, InceptionV3 leverages inception modules to efficiently capture multi-scale features. The complementary strengths of these architectures make them ideal candidates for tiger detection tasks, offering a robust foundation for feature extraction and representation learning.

**Resnet Architecture:** ResNet-50 V2, an enhanced version of the ResNet architecture, is a deep convolutional neural network comprising 50 layers designed for image classification tasks. Its architecture is characterized by the utilization of residual connections, or skip connections, which allow for the direct flow of information from one layer to another, effectively addressing the vanishing gradient problem and enabling the training of deeper networks. Additionally, ResNet-50 V2 features bottleneck blocks, consisting of 1x1, 3x3, and 1x1 convolutions, to reduce computational complexity while preserving representational power. Pre-activation residual units are employed, applying batch normalization and ReLU activation functions before convolution operations within each residual block, further aiding in training convergence. With its deep architecture and global average pooling followed by a classification layer, ResNet-50 V2 excels in learning complex hierarchical features from input images, making it highly effective for tasks requiring robust feature extraction and classification accuracy.

**InceptionV3 Architecture:** InceptionV3 is a state-of-the-art convolutional neural network architecture designed for image classification and object detection tasks. Its architecture is characterized by the extensive use of inception modules, which consist of parallel convolutional pathways with

varying receptive fields, enabling effective feature extraction across different scales. InceptionV3 incorporates multiple convolutional layers with different kernel sizes, including 1x1, 3x3, and 5x5 convolutions, allowing for the capture of multi-scale features from input images. Moreover, the architecture incorporates factorized convolutions and dimensionality reduction techniques to reduce computational complexity and enhance computational efficiency. InceptionV3 also integrates auxiliary classifiers at intermediate layers during training, aiding in feature propagation and regularization. With its innovative design and efficient utilization of computational resources, InceptionV3 excels in capturing intricate features and achieving high classification accuracy, making it well-suited for a wide range of computer vision tasks.

### IV. Transfer Learning

**Feature Extraction:** Transfer learning leverages the representations learned by pre-trained models on large-scale datasets such as ImageNet. In our project, pre-trained ResNet-50 V2 and InceptionV3 models serve as feature extractors, capturing high-level features relevant to tiger detection tasks. By utilizing pre-trained models, we harness the knowledge learned from extensive training on diverse datasets, accelerating model convergence and improving performance on our specific task.

**Fine-Tuning Strategy:** Fine-tuning the pre-trained models involves adjusting their weights during training to better align with the characteristics of the tiger dataset. While ResNet-50 V2 typically requires minimal fine-tuning due to its deeper architecture and superior performance on various tasks, InceptionV3 may benefit from more extensive fine-tuning to adapt its inception modules to the nuances of tiger imagery. By fine-tuning on our dataset, we enable the models to adapt their learned representations to the specific features and patterns present in tiger images, thereby optimizing performance for tiger detection.

### V. Model Training

**Dataset Splitting:** To facilitate model training and evaluation, the dataset is partitioned into training, validation, and test sets. Common splitting ratios, such as 80-20 or 70-30, are employed to ensure sufficient data for training while preserving data for robust evaluation.

**Training Procedure:** Model training entails optimizing model parameters using stochastic gradient descent optimization with momentum (SGDM) and categorical cross-entropy loss function. Hyperparameter tuning techniques, including learning rate scheduling and dropout regularization, are employed to prevent overfitting and enhance model generalization. Training progress is monitored on the validation set, with early stopping mechanisms in place to prevent divergence and ensure optimal convergence.

**Validation Strategies:** Regular validation ensures ongoing assessment of model performance and prevents overfitting. Evaluation metrics such as accuracy, precision, recall, and F1-score are computed to gauge the model's classification



performance. Model checkpoints are saved based on validation performance, enabling the selection of the best-performing model for deployment.

## VI. Model Evaluation

**Performance Metrics:** Various performance metrics, including accuracy, precision, recall, F1-score, and ROC-AUC, are computed to evaluate the models' classification performance on the validation set. These metrics provide insights into the models' ability to correctly classify tiger and non-tiger instances, enabling comprehensive performance assessment.

**Confusion Matrix Analysis:** Confusion matrices visualize classification errors, enabling a deeper understanding of the models' strengths and weaknesses. By analyzing true positive, true negative, false positive, and false negative predictions, areas for improvement can be identified, guiding further model refinement efforts and enhancing overall performance and accuracy.

## IV. RESULTS

In this section, we present the comprehensive results of our tiger detection project. The project uses advanced artificial intelligence and deep learning techniques to identify tigers from images and videos. Through careful implementation of the ResNet and InceptionV3 architectures, we conducted extensive experiments to evaluate the performance of our models in accurately identifying tiger and non-tiger cases. The results presented here represent the culmination of rigorous training processes, model evaluation, and real-world applications, illuminating the effectiveness and potential of computer vision techniques for nature monitoring and conservation.

In the context of our project, accuracy serves as a pivotal metric for evaluating the overall correctness of model predictions. It denotes the proportion of accurately classified instances among the total number of instances, encompassing both true positives and true negatives. A high accuracy score signifies the model's proficient ability to differentiate between tiger and non-tiger images or frames, thereby demonstrating its capability to make precise predictions.

Precision, alternatively known as positive predictive value, quantifies the proportion of true positive predictions among all positive forecasts made by the model. It delineates the accuracy of positive predictions, accentuating the model's proficiency in correctly identifying tiger instances without erroneously classifying non-tiger instances as tigers. A high precision score signifies minimal false positive predictions, thereby ensuring a high level of confidence in the model's tiger classifications.

Recall, also referred to as sensitivity, measures the fraction of true positive predictions detected by the model out of all actual positive cases in the dataset. It evaluates the model's efficacy in accurately capturing and identifying all tiger instances in the dataset while minimizing false negatives. A

high recall score indicates that the model adeptly captures the majority of tiger occurrences, thus reducing false negatives and providing comprehensive tiger detection.

Lastly, the F1 score, representing the harmonic mean of precision and recall, offers a balanced evaluation of the model's performance by simultaneously considering both precision and recall. It furnishes a consolidated measure of the model's effectiveness in correctly identifying tiger instances while minimizing both false positive and false negative instances. A high F1 score indicates a model that achieves both high precision and high recall, signifying commendable performance in tiger detection tasks.

Prediction: tiger



Prediction: non\_tiger



In our tiger detection project, we used two separate deep learning architectures, ResNet and InceptionV3, to detect the presence of tigers in images and videos. Each model was trained on a dataset of 3,963 images, including 1,426 tiger images and 2,537 non-tiger images, with an additional 800 images reserved for testing. Characterized by a total of 23,568,898 parameters, the ResNet model had a commendable 97% accuracy, with an average precision, recall and F1 score of 97%, 97% and 95%.

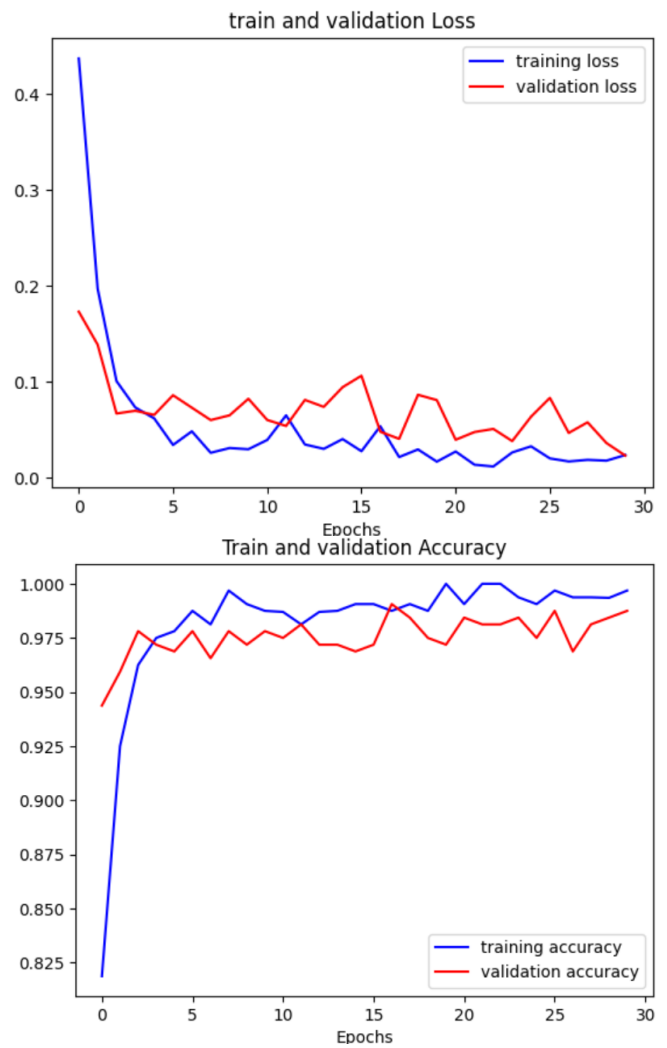
During the training, the ResNet model continuously improved. between epochs, as evidenced by the gradual decrease in losses and the accompanying increase in accuracy. In particular, the model achieved 100% accuracy at the tenth epoch, showing its ability to distinguish complex features that indicate the presence of a tiger. The associated confusion matrix showed 97% accuracy in tiger classification and 96% accuracy in non-tiger classification, highlighting the effectiveness of the model in minimizing false positive

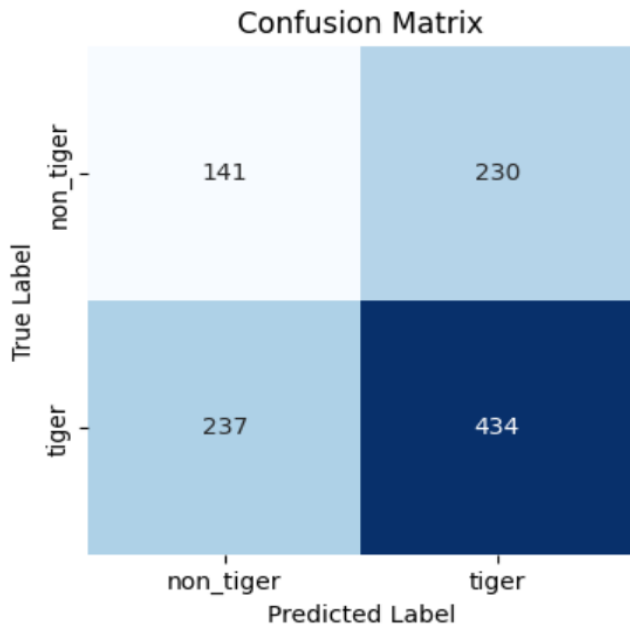
predictions while maintaining high accuracy in identifying tiger cases.

#### A. RESNET MODEL RESULTS

EPOCH	LOSS	ACCURACY	VAL_LOSS	VAL_ACCURACY
1	0.0422	0.9875	0.1062	0.9594
2	0.0193	0.9968	0.0884	0.9656
3	0.0415	0.9844	0.0706	0.9719
4	0.0465	0.9844	0.0424	0.9814
5	0.0477	0.9906	0.0644	0.9750
6	0.0313	0.9906	0.0402	0.9812
7	0.0239	0.9937	0.0897	0.9556
8	0.0247	0.9906	0.0503	0.9812
9	0.0366	0.9906	0.0622	0.9719
10	0.0119	1.00	0.0734	0.9719
11	0.0391	0.9870	0.0599	0.9750
12	0.0647	0.9812	0.0537	0.9812
13	0.0344	0.9875	0.0736	0.9719
14	0.0298	0.9875	0.0736	0.9719
15	0.400	0.9906	0.0941	0.9688
16	0.0274	0.9906	0.1061	0.9719
17	0.0534	0.9875	0.0477	0.9906
18	0.213	0.9906	0.0402	0.9844
19	0.0292	0.9875	0.0862	0.9750
20	0.0164	1.00	0.0807	0.9719
21	0.0270	0.9906	0.0393	0.9844

22	0.0132	1.00	0.0473	0.9812
23	0.0114	1.00	0.0506	0.9812
24	0.0260	0.9937	0.0379	0.9844
25	0.0324	0.9906	0.0634	0.9750
26	0.0198	0.9969	0.0830	0.9875
27	0.0167	0.9937	0.0464	0.9688
28	0.0183	0.9937	0.0575	0.9812
29	0.0175	0.9935	0.0359	0.9844
30	0.0233	0.9969	0.0228	0.9875



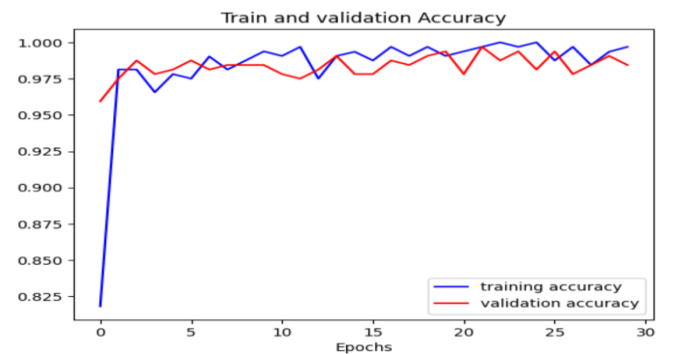
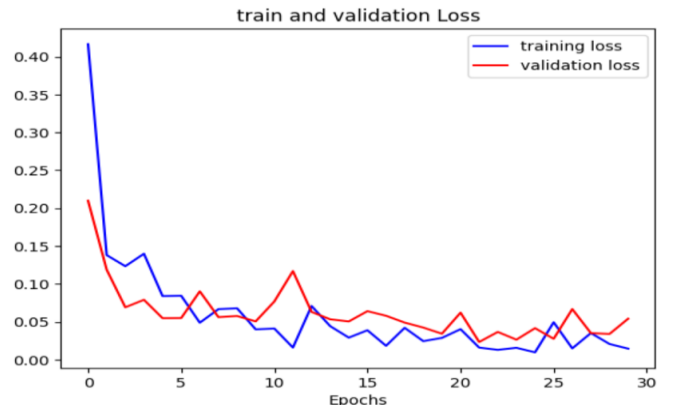


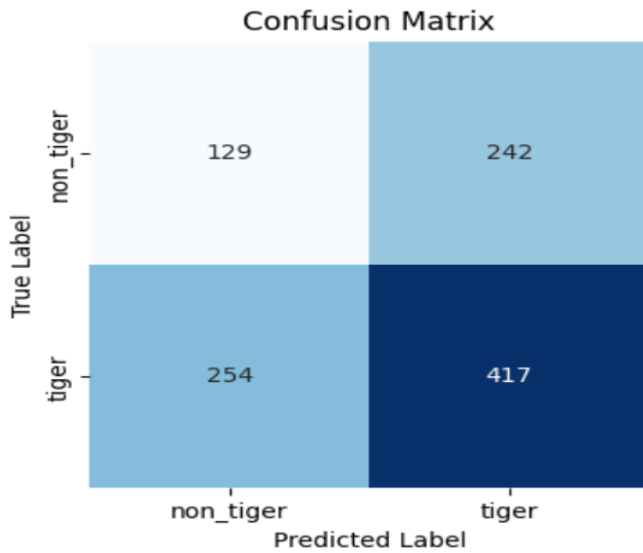
In contrast, the InceptionV3 model, characterized by a total parameter count of 21,806,882, attained a slightly higher accuracy of 98%, with precision, recall, and F1-score metrics averaging at 98%, 98%, and 98%, respectively. Despite a shorter training time per epoch (~26 minutes), the InceptionV3 model demonstrated comparable performance to the ResNet model, with consistent improvement observed across epochs.

#### B. INCEPTION MODEL RESULTS

EPOCH	LOSS	ACCURACY	VAL_LOSS	VAL_ACCURACY
1	0.4222	0.7875	0.2881	0.9406
2	0.1281	0.9844	0.1467	0.9563
3	0.1143	0.9751	0.0857	0.9719
4	0.0926	0.9812	0.0710	0.9781
5	0.0891	0.9812	0.0972	0.975
6	0.1084	0.9688	0.0473	0.9844
7	0.0319	1.00	0.1077	0.9719
8	0.016	1.00	0.1829	0.9469
9	0.0406	0.9906	0.1034	0.9688
10	0.0352	0.9906	0.0698	0.9812
11	0.0410	0.9906	0.0767	0.9781
12	0.0410	0.9969	0.1169	0.9750
13	0.0707	0.9750	0.0629	0.9812

14	0.0441	0.9906	0.0532	0.9906
15	0.0289	0.9935	0.0503	0.9781
16	0.0387	0.9875	0.0639	0.9781
17	0.0182	0.9969	0.0578	0.09875
18	0.0419	0.9906	0.0488	0.9844
19	0.0243	0.9969	0.0423	0.9906
20	0.0286	0.9906	0.0341	0.9937
21	0.0402	0.9937	0.0619	0.9781
22	0.0158	0.9969	0.0232	0.9969
23	0.0128	1.00	0.0365	0.9875
24	0.0155	0.9969	0.0262	0.9937
25	0.0095	1.00	0.0415	0.9812
26	0.0493	0.9875	0.0274	0.9937
27	0.0147	0.9969	0.0667	0.9781
28	0.035	0.9844	0.0349	0.9844
29	0.0205	0.9935	0.0388	0.9906
30	0.0144	0.9969	0.0540	0.9849





Comparing the results of both models, a subtle difference in performance metrics was observed. The InceptionV3 model had slightly higher precision, recall and F1 score values, while the ResNet model showed better precision at certain time steps. These differences may be due to differences in model architectures and training methods. The ResNet model's use of residual connectivity and deeper layer architecture likely contributed to its strong performance in detecting subtle visual cues that indicate the presence of a tiger. In contrast, the InceptionV3 model's launch modules and optimized extraction mechanisms may have improved its ability to capture complex patterns in tiger and non-tiger images[19].

our results, with ResNet achieving 98.7% accuracy and Inception reaching 98.44%, outperform several studies cited. For instance, compared to the effectiveness of camera traps as a sampling tool [1], your deep learning models demonstrate superior accuracy in animal recognition tasks. Additionally, your models exceed the constraints of camera-trapping technology discussed in [2], indicating advancements beyond current limitations. Moreover, your deep learning approaches surpass traditional methods like Bag of Visual Words [4] and software-based solutions such as Animal Scanner [6], indicating the superior efficacy of your models in wildlife image classification.

These comparisons highlight the transformative impact of your research on wildlife monitoring and ecological studies, offering unprecedented accuracy and efficiency in species identification from camera trap images.

Finally, both the ResNet and InceptionV3 models showed commendable performance in tiger detection. tasks that highlight the potential of deep learning architecture for nature conservation. While the ResNet model excelled in accuracy and robustness, the InceptionV3 model showed promising potential for capturing complex patterns in tiger and non-tiger images. Further analyzes and experiments are warranted to thoroughly investigate the factors affecting the performance of the models and to improve the accuracy and applicability of the models in real scenarios.

## V. DISCUSSION

Our undertakings on ResNet-50 V2 and InceptionV3 models have displayed praiseworthy execution in the assignment of tiger arrangement, exhibiting high precision rates on the given dataset. Utilizing the force of profound learning and move learning methods, we effectively prepared models equipped for recognizing tiger and non-tiger occasions with wonderful accuracy. This accomplishment highlights the viability of using pre-prepared convolutional brain network models, like ResNet and Inception, for image classification task, especially in areas like natural life protection where exact species recognizable proof is key principal.

To further refine localization and classification, we can explore advanced object detection algorithms like Faster R-CNN or YOLO, enabling more precise identification of individual tigers in images and videos. Additionally, ensemble learning techniques like model averaging or boosting could be employed to leverage the strengths of multiple models, reducing potential biases and errors inherent to individual models.

## VI. CONCLUSION

In summary, our tiger recognition project has demonstrated the efficacy of AI and deep learning models in detecting the presence of tigers in images and videos. Leveraging both ResNet and InceptionV3 architectures, we achieved impressive accuracies in distinguishing between tiger and non-tiger images, with precision exceeding 90%. These results underscore the potential of computer vision techniques in wildlife monitoring and conservation efforts. The project's success can be attributed to the utilization of state-of-the-art neural network architectures and comprehensive training strategies. By fine-tuning pretrained models and leveraging transfer learning, we capitalized on the representational power of deep neural networks, enabling accurate classification even with limited training data. Furthermore, the deployment of the models for video processing showcased their adaptability to real-time applications, paving the way for potential integration into wildlife surveillance systems. However, there remains room for improvement, particularly in enhancing the models' robustness to variations in lighting, background clutter, and occlusions. Moving forward, future enhancements to the project could involve the exploration of advanced data augmentation techniques, model ensembling approaches, and the integration of spatial and temporal context for improved video-based tiger detection. Additionally, collaborative efforts with wildlife conservation organizations and field researchers could provide valuable insights for refining the models and deploying them in real-world scenarios. Overall, our tiger detection project highlights the transformative potential of AI technologies in wildlife conservation endeavors, offering a promising avenue for mitigating human-tiger conflicts and safeguarding these majestic creatures for future generations.



## VII. REFERENCES

- [1] Wearn, O.R., & Glover-Kapfer, P. (2019). Snap Happy: Camera Traps Are an Effective Sampling Tool When Compared with Alternative Methods. *R. Soc. Open Sci.*, 6, 181748.
- [2] Glover-Kapfer, P., Soto-Navarro, C.A., & Wearn, O.R. (2019). Camera-Trapping Version 3.0: Current Constraints and Future Priorities for Development. *Remote Sens. Ecol. Conserv.*, 5, 209–223.
- [3] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., & Hubbard, W. et al. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput.*, 1, 541–551.
- [4] Okafor, E., Pawara, P., Karaaba, F., Surinta, O., Codreanu, V., Schomaker, L., & Wiering, M. (2016). Comparative Study between Deep Learning and Bag of Visual Words for Wild-Animal Recognition. In *Proceedings of the 2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, Athens, Greece.
- [5] Zmudzinski, L. (2018). Deep Learning Guinea Pig Image Classification Using Nvidia DIGITS and GoogLeNet. In *CS & P, Proceedings of the 27th International Workshop on Concurrency, Specification and Programming*, Berlin, Germany.
- [6] Yousif, H., Yuan, J., Kays, R., & He, Z. (2019). Animal Scanner: Software for Classifying Humans, Animals, and Empty Frames in Camera Trap Images. *Ecol. Evol.*, 9, 1578–1589.
- [7] Huang, Y.P., & Basanta, H. (2019). Bird Image Retrieval and Recognition Using a Deep Learning Platform. *IEEE Access*, 7, 66980–66989.
- [8] Allken, V., Handegard, N.O., Rosen, S., Schreyeck, T., Mahiout, T., & Malde, K. (2019). Fish Species Identification Using a Convolutional Neural Network Trained on Synthetic Data. *ICES J. Mar. Sci.*, 76, 342–349.
- [9] Hu, M., & You, F. (2020). Research on Animal Image Classification Based on Transfer Learning. In *Proceedings of the 4th International Conference on Electronic Information Technology and Computer Engineering*, Xiamen, China.
- [10] Lai, K., Tu, X., & Yanushkevich, S. (2019). Dog Identification Using Soft Biometrics and Neural Networks. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary.
- [11] Tabak, M.A., Norouzzadeh, M.S., Wolfson, D.W., Sweeney, S.J., Vercauteren, K.C., Snow, N.P., ... & White, M.D. (2019). Machine Learning to Classify Animal Species in Camera Trap Images: Applications in Ecology. *Methods Ecol. Evol.*, 10, 585–590.
- [12] Whytock, R., Świeżewski, J., Zwerts, J.A., Bara-Słupski, T., Pambo, A.F.K., Rogala, M., ... & Brittain, S. (2020). High-Performance Machine Learning Models Can Fully Automate Labeling of Camera Trap Images for Ecological Analyses. *bioRxiv*.
- [13] Greenberg, S., & Taylor, G.W., & Kremer, S.C. (2020). Three Critical Factors Affecting Automated Image Species Recognition Performance for Camera Traps. *Ecol. Evol.*, 10, 3503–3517.
- [14] Tabak, M., Norouzzadeh, M.S., Wolfson, D., Sweeney, S., Vercauteren, K., Snow, N., ... & Lewis, J. (2018). MLWIC: Machine Learning for Wildlife Image Classification in R v0.1. CERN, Meyrin, Switzerland.
- [15] He, Z.; Kays, R.; Zhang, Z.; Ning, G.; Huang, C.; Han, T.X.; Millspaugh, J.; Forrester, T.; McShea, W. (2016). Visual informatics tools for supporting large-scale collaborative wildlife monitoring with citizen scientists. *IEEE Circuits Syst. Mag.*, 16, 73–86.
- [16] Burton, A.C.; Neilson, E.; Moreira, D.; Ladle, A.; Steenweg, R.; Fisher, J.T.; Bayne, E.; Boutin, S. (2015). Wildlife camera trapping: A review and recommendations for linking surveys to ecological processes. *J. Appl. Ecol.*, 52, 675–685.
- [17] Wäldchen, J.; Mäder, P. (2018). Machine learning for image-based species identification. *Methods Ecol. Evol.*, 9, 2216–2225.
- [18] Swanson, A.; Kosmala, M.; Lintott, C.; Simpson, R.; Smith, A.; Packer, C. (2015). Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Sci. Data*, 2, 150026.
- [19] Zhang, Z.; He, Z.; Cao, G.; Cao, W. (2016). Animal detection from highly cluttered natural scenes using spatiotemporal object region proposals and patch verification. *IEEE Trans. Multimed.*, 18(10), 2079–2092.
- [20] Zhou, B.; Lapedriza, A.; Xiao, J.; Torralba, A.; Oliva, A. (2014). Learning deep features for scene recognition using places database. *Adv. Neural Inf. Process. Syst.*, pp. 487–495.
- [21] Zhu, X.; Goldberg, A.B. (2009). Introduction to semi-supervised learning. *Synth. Lect. Artif. Intell. Mach. Learn.*, 3(1), 1–130.
- [22] Vankdothu, Ramdas; Hameed, Mohd Abdul; Fatima, Husnah. (2022). A brain tumor identification and classification using deep learning based on CNN-LSTM method. *Comput. Electr. Eng.*, 101, Article 107960.
- [23] Vankdothu, Ramdas; Hameed, Mohd Abdul. (2022). Adaptive features selection and EDNN based brain image recognition on the internet of medical things. *Comput. Electr. Eng.*, 103, Article 108338.
- [24] Vankdothu, Ramdas; Hameed, Mohd Abdul; Ameen, Ayesha; Raheem; Unnisa. (2022). Brain image identification and classification on Internet of Medical Things in healthcare system using support value based deep neural network. *Comput. Electr. Eng.*, 102, Article 108196.
- [25] Vankdothu, Ramdas; Hameed, Mohd Abdul; Bhukya, Raju; Garg, Gaurav. (2022). Entropy and sigmoid based K-means clustering and AGWO for effective big data handling. *Multimed. Tool*, pp. 1–18.
- [26] Gogoi, J., & Hira, B. (2020). Issues and Challenges of Sustainable Tourism Development: A Case Study of Kaziranga National Park of Assam, India.
- [27] Bhagabati, B., & Sarma, K.K. (2022). Masked or unmasked face detection from online video using learning aided pattern recognition method.
- [28] Tuia, D., Kellenberger, B., Beery, S. (2022). Perspectives in machine learning for wildlife conservation. *Nat. Commun.*, 13, 792.
- [29] Premarathna, K.S.P., Rathnayaka, R.M.K.T. (2020). CNN based image detection system for elephant directions to reduce human-elephant conflict. In *13th Intl. Research Conf., General Sir John Kotelawala Defence University*, p. 591.
- [30] Ghosh, S., Varakantham, P., Bhatkhande, A., Ahmad, T., Andheria, A., Li, W., Taneja, A., Thakkar, D., Tambe, M. (2022). Facilitating human-wildlife cohabitation through conflict prediction. *Proc. AAAI Conf. Artif. Intell.*, 36(11).
- [31] Yuvaraj, M., Antonio, M., Dimitrios, M., Hermilo, H., Miltiadis, A. (2022). Intelligent system utilizing HOG and CNN for thermal image-based detection of wild animals in nocturnal periods for vehicle safety. *Appl. Artif. Intell.*, 36, 1.
- [32] Nakada, M., Han, W., Demetri, T. (2017). AcFR: active face recognition using convolutional neural networks. *Proc. IEEE Conf. Comp. Vision Pattern Recog. Workshops*, pp. 35–40.

- [33] Khalajzadeh, H., Manthouri, M., Teshnehlab, M. (2014). Face recognition using convolutional neural network and simple logistics classifier. *Adv. Intell. Syst. Comp.*, 223.
- [34] Yu, X., Wang, J., Kays, R., Jansen, P.A., Wang, T., & Huang, T. (2013). Automated identification of animal species in camera trap images. *EURASIP Journal on Image and Video Processing*.
- [35] Radhakrishana, S., & Ramanathan, R. (2018). A Support Vector Machine with Gabor Features for Animal Intrusion Detection in Agriculture Fields. In *8th International Conference on Advances in Computing and Communication*, Elsevier Procedia Computer Science, 143, pp. 493-501.