International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3303-3308
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0423

Automated Tiger Detection System Using Faster R-CNN and MobileNet Architecture

Khemnar Kavita Changdeo[1], *Dr. Monika Rokade*[2], *Dr. Sunil Khatal*[3]
[1]*Student, Dept. of Computer Engg. Sharadchandra Pawar College of Engineering, Otur, India*
[2,3]*Prof, Dept of Computer Engg. Sharadchandra Pawar College of Engineering, Otur, India*
Email ID: *kavitakhemnar1@gmail.com*[1], *monikarokade4@gmail.com*[2], *khatalsunils88@gmail.com*[3]

Abstract
The growing importance of wildlife conservation is highlighted by the declining numbers of endangered species like the tiger. It is essential that we develop tools for the precise and automated monitoring of tiger populations in the wild for conservation purposes. Current techniques for population monitoring rely on laborious manual checks, and the use of camera traps, which are susceptible to random human error. The increased use of deep learning techniques has led to increased interest in the automation of animal detection through the use of object detection models. Unfortunately, in the case of tiger detection, there are a number of unique challenges, such as occlusion, changes in the light, the presence of dense undergrowth, and the small size of the target. We propose a detection system that is efficient in terms of tiger detection and based on deep learning techniques, particularly the Faster R-CNN model with a MobileNet backbone. MobileNet supports a lightweight architecture for real-time use, while Faster R-CNN is able to achieve a high detection rate because of its regional proposal approach. These values represent the high levels of diverse safety detection performance cited in the methodology. In closing, the proposed Faster R-CNN with MobileNet backbone architecture works as a strong detector and value-adding MobileNet backbone architecture. The system proposed has a place in the systems used in the monitoring of wildlife, providing the value of detection systems in the field for tracking and in the protection of wildlife.

Keywords: *Deep Learning; Faster R-CNN; MobileNet; Object Detection; Tiger Detection; Wildlife Monitoring.*

1. Introduction

The rapid loss in biodiversity has sparked a worldwide concern for wildlife conservation. Tigers, in particular, are threatened by habitat loss, poaching, and conflicts with humans, and are among the most endangered species. Understanding their population for monitoring tiger population is a primary factor for the development of tiger population conservation strategies. The population dynamics of tigers in the wild is a complex and inefficient system for the large-scale use of delayed conservation techniques and the collection of data necessary to understand the population dynamics of tigers in the wild to the manual survey and camera trap analysis. The complexity of detecting wildlife has been emphasized in case of long-tail object detection problems in [1] and [7]. Due to class imbalance, there are significantly less tiger images in comparison to other objects. Although newer models based on the YOLO architecture [2], [3], [11], and [12] have been able to increase speed of detection, they tend to lose accuracy in more difficult, long-tail detection situations. For instance, detecting less frequent tigers in objects present in the training data set. Furthermore, while models utilizing the transformer architecture [4], [5], [13] have shown to have high potential, the high computational power required has limited their use in real-time detection applications where fast responses are required. The detection of small, distant objects, and tigers in forests or indistinct images, is problematic as discussed in [8], [9], and [10]. To tackle this, feature pyramid networks and attention mechanisms have been used. This paper presents a tiger detection model based on Faster R-CNN using MobileNet as a backbone. The MobileNet architecture promotes detection on resource-constrained platforms (edge devices), including drones and special-purpose IoT devices for wildlife monitoring. The Faster R-CNN structure

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3303-3308
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0423

increases detection accuracy by proposing candidate bounding boxes (regions of interest) and iteratively improving those boxes through classification and regression. The system is trained on a dataset sourced and annotated on the Roboflow platform. The model is fine-tuned with the transfer learning paradigm, data augmentation, and hyperparameter optimization. The goal is to maximize detection accuracy and system efficiency. This study adds to the body of knowledge in automated and scalable wildlife monitoring solutions by conducting automated tiger detection. The solution outlines methods to address challenges posed by occlusions, class imbalance, and variable environments to promote real-world applicability.

2. Literature Survey

Ye et al. [1] introduced a method to enhance the calibration of margins and average precision losses to improve detection of long-tailed objects. Their method goes about resolving the class imbalance problem by modifying the loss function to emphasize focus on losses from minor classes. This method works for wildlife detection when measuring instances of tigers which are few in contrast to the backdrop data. The study revealed that detection precision improved on the data sets with a lack of balance. The study demonstrates the focus and significance of loss and deep learning model optimization. Gao et al. [2] came up with the model YOLO-Parallel which seeks to advance the construction of gradient modeling for long-tailed object detection. Their method seeks to improve the representation of features through positive gradient flow, which in turn improves detection in complicated zones. The model, however, focuses on the quickness of function rather than the optimization of function. It also offers an overarching technique to navigate through imbalanced data sets. Jocher et al [3] created the YOLOv8 model which is the best of its age for objects detection in real time. Compared to other versions from the YOLO family, the model has a better speed and accuracy balance. It also has new and better optimized model architecture and training strategies. Although YOLOv8 is an improved model, it will not accurately detect small and occluded objects which are a key component in wildlife conservation monitoring. It is a YOLOv8 drawback. Pan et al. [4] examined how self-attention mechanisms can shed light on the workings of convolutional neural networks. Their work shows how attention and convolution can be integrated to refine the feature extraction process. This method is useful to improve the detection of objects in complicated scenarios. On the downside, the method increases the computational requirements. The study is an initial step towards the creation of hybrid deep learning models. Tu et al. [5] proposed MaxViT and multi-axis vision transformers that depict local and global features optimization providing accuracy for image recognition and object detection. However, the computational cost is considerably greater than CNN-based models, making it less practical for real-time systems for wildlife monitoring. Rajagopal et al. [6] proposed a Convolutional Gated MLP model, which integrates convolution with MLP structures. This model promotes better feature representation and learning, which is great for classification, but it is not as refined for object detection frameworks, supporting the notion of combining different profiles of deep learning. Li et al. [7] present a Gaussian Clouded Logit Adjustment as a technique for long-tailed visual recognition. This technique proposes to mitigate the balance of classification problem by adjusting the logit distribution and it works particularly well for the class imbalance data, which is the case with a small number of sample tigers. The findings confirm that proper data distribution is a fundamental aspect of the problem. Fujii et al. [8] discussed remote bird detection in terms of drone safety. They have created a framework and dataset for detecting small objects. The authors describe issues related to low resolution and background interference. These issues are analogous to those involved in detecting tigers in a forest. The authors provide methods and solutions to these problems, and improve small object detection. Kondo et al. [9] created a bird detection challenge for small object detection. He also created the dataset, evaluation criteria, and detection method benchmarks. He described the problems related to the detection of small objects at long distances. He also described the problems associated with the lack of evaluation methods. His findings relate to the problems of systems for detecting wildlife.

Sun et al. [10] created a dataset for detecting birds in

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3303-3308
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0423

video surveillance. Their dataset contains a variety of environments and scales of objects. The authors describe the problems related to the detection of motion blur and occlusion. Their findings can be extrapolated to detecting tigers in varying environments.

3. Proposed Methodology

A faster R-CNN with a MobileNet backbone is being implemented as the deep learning object detection framework for the proposed tiger detection system. The tiger detection system automatically identifies tigers in pictures taken within the tiger's natural setting. The proposed methodology involves stages such as the establishment of a data acquisition protocol, a data preprocessing protocol, a data feature extraction protocol, a data region proposal, a data object classification protocol, and evaluation of system performance. Real-world challenges, including occlusion, varying lighting, clutter, and small size challenges, are addressed. MobileNet's lightweight backbone enables the system's computational efficiency. The detection accuracy is increased as a result of the two-stage detection proposed by the framework. Region of Interest (RoI) classification is followed by the proposal of regions

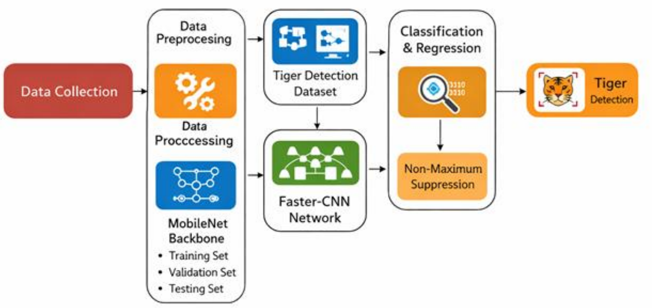

Figure 1. **System Architecture of Proposed System**

List of Modules and Functionality

Data Collection Module: This module collects images of tigers from the Roboflow dataset. It collects images from different regions, such as, forests with different lighting and tigers in different positions. Good data collection increases the accuracy of the model. The collected data is processed and organized for further operations.

Data Preprocessing Module: The preprocessing module works with the raw images that the model will train on. It resizes the images to one dimension and normalizes the pixel size. Generalization is improved by the method of data augmentation, by applying it in the combination of: flipping, rotating, and adjusting the brightness. This module changes the annotations to a format that is compatible with the Faster R-CNN.

Dataset Splitting Module: This module separates the dataset into 3 groups: one for training, another for validation, and the last for testing. It makes sure the images of tigers in the dataset are equally divided in all of the groups. Proper division helps the model evaluation to be unbiased. Also, it prevents the model from being overfitted by conducting a validation on data that is not previously seen to the model.

Feature Extraction Module (MobileNet Backbone): The MobileNet feature extraction module captures key features from the input images. It uses depthwise separable convolutions to lessen the load on the system. The module captures and generates feature maps that demonstrate both the spatial and the semantic. These maps are to be used in the detection of objects in the subsequent phases.

Model Training Module: This module is in charge of training the Faster R-CNN model. It incorporates transfer learning and the use of pre-trained weights, positively impacting performance. The model training employs certain optimization strategies, Adam optimizer being one of them. Training is performed on a number of epochs in hopes of refining the accuracy.

Model Testing Module: The final module to partake in exporting the system's final output. The images are encapsulated with bounding boxes enclosing the Tigers that have been successfully detected, and the confidence scores are documented for each detection. The output is of use for both monitoring as well as analyses

Evaluation Module: This module is reserved for the evaluation of the model's performance. The evaluated parameters are Precision, Recall, F1 Score, and IoU. With the evaluation of these parameters, the effectiveness of the detection system and the reliability of the model are evaluated the module ends up speaking for the model's reliability.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3303-3308
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0423

Algorithm: Tiger Detection Faster RCNN

Input: Image dataset D (train, validation, test), Pre-trained MobileNet weights

Output: Detected tiger with bounding boxes and confidence score

Step1: Load dataset D and split into Train, Validation, and Test sets

Step 2: FOR each image in D DO
Resize image to 640x640
Normalize pixel values
Apply data augmentation (flip, rotate, brightness)
END FOR

Step 3: Initialize Faster R-CNN with MobileNet backbone and RPN

Step 4: FOR epoch = 1 to N DO
FOR each training image I DO
Extract feature maps using MobileNet
Generate region proposals using RPN
Apply ROI Pooling
Predict class (Tiger/Background) and bounding boxes
Compute total loss (classification + localization)
Update model weights
END FOR
Validate model using validation dataset
END FOR

Step 5: FOR each test image T DO
Extract features using MobileNet
Generate region proposals using RPN
Apply ROI Pooling
Predict class and bounding boxes
Apply Non-Maximum Suppression (NMS)
Output detected tiger
END FOR

Step 6: Compute evaluation metrics (Precision, Recall, F1 Score, IoU)

Step 7: END

4. Results And Discussion

The detection system for tigers in this proposal uses a Faster R-CNN model with a MobileNet Backbone, and is subjected to evaluation based on conventional criteria. The model training consisted of 2,500 images for training, 500 images for validation, and 500 images for testing. This detection system demonstrates commendable performance in identifying tigers in divergent terrains such as dense forests, low and high light conditions, and situations where the subject is partially obscured.

The results show the robustness and accuracy of the proposed model for the tiger detection task, as well as the ability to outperform Faster R-CNN single-stage detectors by achieving better localization, as demonstrated by the better IoU score. MobileNet, therefore, allows for substantial reduction of the computational cost without sacrificing detection ability.

A. Analysis of Training Performance

Figure 2 the training loss curve appearing plummeting consistently over time over the increment of the epoch can be attributed to the model learning fundamental features. When training began, loss was drastically high around (approx.) 0.28, however, in the long run, after 60 epochs, loss stabilizes at roughly 0.025. Such loss stabilization trends are synonymous with effective learning and, model convergence with regards to over fitting. Furthermore, smooth descent of the curve could indicate the rate and strategy of optimization coupled with learning are reasonable.

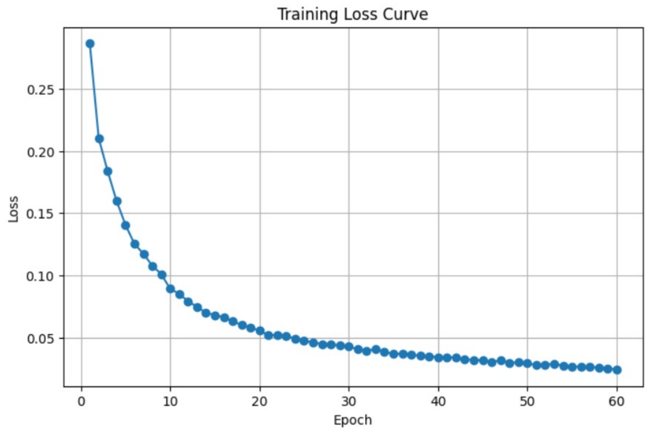

FIGURE 2. **Training Loss Curve**

B. Results of Detection

The figure 3 system detects the presence of a tiger in the test images and demonstrates high levels of confidence. Predicted images contain a bounding box that \encapsulates a tiger with a confidence level of 1.00. In particular, the model was able to thrive in test images that was accompanied with a dense and complex background (vegetation). Such scenarios display the prowess of the Faster R-CNN framework in combination with MobileNet

IRJAEH
International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 05 May 2026
Page No: 3303-3308
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0423

in features extraction.

FIGURE 3. **Tiger Results of Detection**

C. Evaluation

The metrics for the new model are;

Table 2: Performance Evaluation Results

Metric	Value
Precision	0.9577
Recall	0.9027
F1 Score	0.9294
Mean IoU	0.8541

The table 2 model has a very high precision which means the number of false positives is minimal. Most of the instances of tigers are detected which is what the recall value indicates. The F1 score shows an optimal balance between the precision and recall of the model. The tigers were accurately localized which is what the high value of the Mean IoU shows**.**

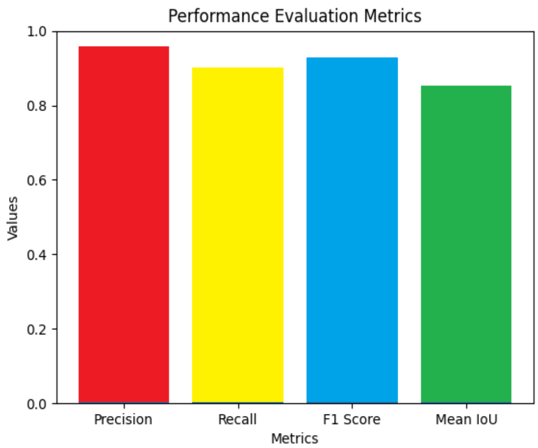

Figure 4. **Performance Evaluation Metrics**

The Figure 4 model's performance in detecting tigers is measured through four major evaluation criteria: Mean Intersection Over Union, F1 Score, Recall, and Precision. The Mean IoU is represented by the most recent data (shown in the graph) at (0.8541) and indicates that there is considerable agreement between the predictor boxes and the actual boxes, and therefore, the model is successful in identifying the objects. Out of all the evaluation metrics, our model performed the best in Precision (0.9577) and the worst in Recall (0.9027). Given that there were hardly any false positive identifications, this positive performance in Precision is related to the very few tiger instances missing (some were likely missed). The high F1 Score (0.9294) indicates an excellent balance between Precision and Recall, meaning the model is very accurate and very complete in its identifications**.**

Conclusion

This research has developed and assessed an effective deep learning-based detection system for tigers using Faster R-CNN with a MobileNet backbone. Proposed Methodology has created a systematic solution for vital issues in wildlife monitoring varying environmental conditions, occlusions, and difficult backgrounds. With Faster R-CNN offering precise object detection and MobileNet promoting effective processing, this system combines detection and processing power.

The Model's dataset structure, comprised of training, validation and testing images, facilitated effective training and evaluation. 0.9577 precision, 0.9027 recall, 0.9294 F1 score and 0.8541 mean IoU showcase the system's accuracy and ability to identify and locate tigers in various conditions. With high precision, false detection is low, and the system demonstrated high recall and is successful in detection of the majority of tiger instances

Journal reference style:

[1]. Ye, Y.; Zhang, T.; Lu, R. Margin and Average Precision Loss Calibration for Long- ail Object Detection. In Proceedings of the 2024 9th International Conference on Computer and Communication Systems (ICCCS), Xi'an, China, 19–22 April 2024; pp. 26–32.
[2]. Gao, X.; Zhao, D.; Yuan, Z. YOLO-Parallel: Positive Gradient Modeling for Long-Tail

Remote Sensing Object Detection. In Proceedings of the IEEE Geoscience Remote Sensing Letters, Athens, Greece, 7–12 July 2024.
[3]. Jocher, G.; Chaurasia, A.; Qiu, J. Ultralytics YOLOv8. 2023. Available online: https://docs.ultralytics.com/zh/models/yolov8/ (accessed on 8 November 2022).
[4]. Pan, X.; Ge, C.; Lu, R.; Song, S.; Chen, G.; Huang, Z.; Huang, G. On the integration of self-attention and convolution. In Proceedings of the IEEE/CVF Conference On Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 815–825.
[5]. Tu, Z.; Talebi, H.; Zhang, H.; Yang, F.; Milanfar, P.; Bovik, A.; Li, Y. Maxvit: Multiaxis vision transformer. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 459–479.
[6]. Rajagopal, A.; Nirmala, V. Convolutional Gated MLP: Combining Convolutions and gMLP. In Proceedings of the International Conference on Big Data, Machine Learning, and Applications, Kenitra, Morocco, 5–6 June 2021; pp. 721–735.
[7]. Li, M.; Cheung, Y.-m.; Lu, Y. Long-tailed visual recognition via gaussian clouded logit adjustment. In Proceedings of the IEEE/CVF Conference On Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 6929–6938.
[8]. Fujii, S.; Akita, K.; Ukita, N. Distant bird detection for safe drone flight and its dataset. In Proceedings of the 2021 17th International Conference on Machine Vision and Applications (MVA), Singapore, 20–22 February 2021; pp. 1–5.
[9]. Kondo, Y.; Ukita, N.; Yamaguchi, T.; Hou, H.-Y.; Shen, M.-Y.; Hsu, C.-C.; Huang, E.-M.; Huang, Y.-C.; Xia, Y.-C.; Wang, C.-Y. Mva2023 small object detection challenge for spotting birds: Dataset, methods, and results. In Proceedings of the 2023 18th International Conference on Machine Vision and Applications (MVA), Singapore, 10–12 March 2023; pp. 1–11.
[10]. Sun, Z.-W.; Hua, Z.-X.; Li, H.-C.; Qi, Z.-P.; Li, X.; Li, Y.; Zhang, J.-C. FBDSV- 2024: Flying Bird Object Detection Dataset in Surveillance Video. arXiv 2024, arXiv:00317.
[11]. Contributors, Y. You Only Look Once Version 5. 2021. Available online: https://github.com/ultralytics/yolov5/ (accessed on 7 October 2022).
[12]. Li, C.; Li, L.; Jiang, H.; Weng, K.; Geng, Y.; Li, L.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W. YOLOv6: A single-stage object detection framework for industrial applications. arXiv 2022, arXiv:02976.
[13]. Wang, C.-Y.; Yeh, I.-H.; Liao, H.-Y.M. Yolov9: Learning what you want to learn using programmable gradient information. arXiv 2024, arXiv:13616.
[14]. Wang, W.; Xie, E.; Li, X.; Fan, D.-P.; Song, K.; Liang, D.; Lu, T.; Luo, P.; Shao, L. Pyramid vision transformer: A versatile backbone for dense prediction without convolutions. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 11–17 October 2021; pp. 568–578.
[15]. Zhang, C.; Chen, Y.; Hao, Z.; Gao, X. An efficient time-domain end-to-end singlechannel bird sound separation network. Animals 2022, 12, 3117.