

# Transfer learning on animal species recognition tasks

*Haotian Zhu*

Imperial College London, Faculty of Natural Sciences, Level 3, Faculty Building, London, SW7 2AZ, UK

Hz1321@ic.ac.uk

---

**Abstract.** In the field of wildlife conservation, it is challenging to distinguish visually similar species, such as leopards, lions and tigers. This poses huge obstacles for automated detection systems. In this study, transfer learning uses YOLOv5 model from Ultralytics repository to enhance the accuracy of the detection. We pre-trained the model on datasets specific to each species and then applied transfer learning across different species pairs (e.g., leopard to lion, lion to tiger). Our results indicate that while models pre-trained on individual species achieved high detection accuracy (mAP@0.5 > 0.95), the effectiveness of transfer learning varied significantly depending on the visual similarity between species. For instance, transferring from leopard to lion demonstrated strong performance (peak mAP@0.5 of 0.91), while transferring from leopard to tiger resulted in lower accuracy (peak mAP@0.5 of 0.80), showing the limitations of transferring learned features between more visually distinct species. These findings indicate that transfer learning largely reduces training time and improves the adaptability of the models when applied to closely related species. However, further improvements are still needed for more distant species pairs and for too small datasets.

**Keywords:** deep learning, transfer learning, object detection, ecology

---

## 1. Introduction

Recent publications and findings in deep learning have significantly improved the accuracy and efficiency of animal detection and classification [1]. Traditional methods, like histogram-based techniques and edge detectors, often become incapable of handling the complexities of real-world animal images. This is caused by varying lighting conditions, backgrounds, and object orientations of these animal images [2]. However, deep learning methods seem to be the perfect solution for this complex task. Among deep learning methods, convolutional neural networks (CNNs) are regarded to be highly effective for such tasks because of their ability to learn complicated features from raw data.

Object recognition in wild animals has been trained using a variety of CNN architectures, including Faster R-CNN, SSD (Single Shot Multibox Detector), and YOLO (You Only Look Once). These architectures locate the animals in images and classify them into different species. However, challenges still remain when trying to differentiate visually similar species, such as lions, tigers, and leopards. This happens more frequently especially when the images are taken in similar environments, with overlapping visual features or these species co-existence with each other.

Transfer learning is a widely used technique in the Deep Learning field. It can be applied to various domains when classifying images [3]. For example, medical imaging and autonomous driving are the two most important fields [4]. The foundation of transfer learning is to use a pre-trained model in the first place, and then fine-tune it on a smaller dataset specific to the task at hand. Utilising transfer learning provides better and faster results compared to training from scratch.

In this study, transfer learning is employed to adapt the knowledge learned from detecting one animal species (e.g., lions) to detect another similar species (e.g., tigers or leopards). This aims to improve the detection performance in situations where labelled data is limited.

Through this study, I provide an efficient solution for accelerating the model training of animal recognition and offer new perspectives on the use of transfer learning between similar species [5]. This improvement accelerates the model training process with transfer learning in animal recognition tasks.

## 2. Background and motivation

Nowadays, the fragmentation and loss of natural habitats along with the increase in human activities are becoming a more and more serious issue around the world in all kinds of ecosystems [6]. The protection of wildlife and their ecosystems is becoming

crucial to large feline animals such as lions, leopards, and tigers. In most of cases, they are the top predators in their respective ecosystems and play an important role in maintaining ecological interaction networks. Unfortunately, these species are more susceptible to these issues, such as poaching, habitat loss, and human-wildlife conflict. As a result, real-time monitoring and accurate identification of these animals are essential to conservation efforts.

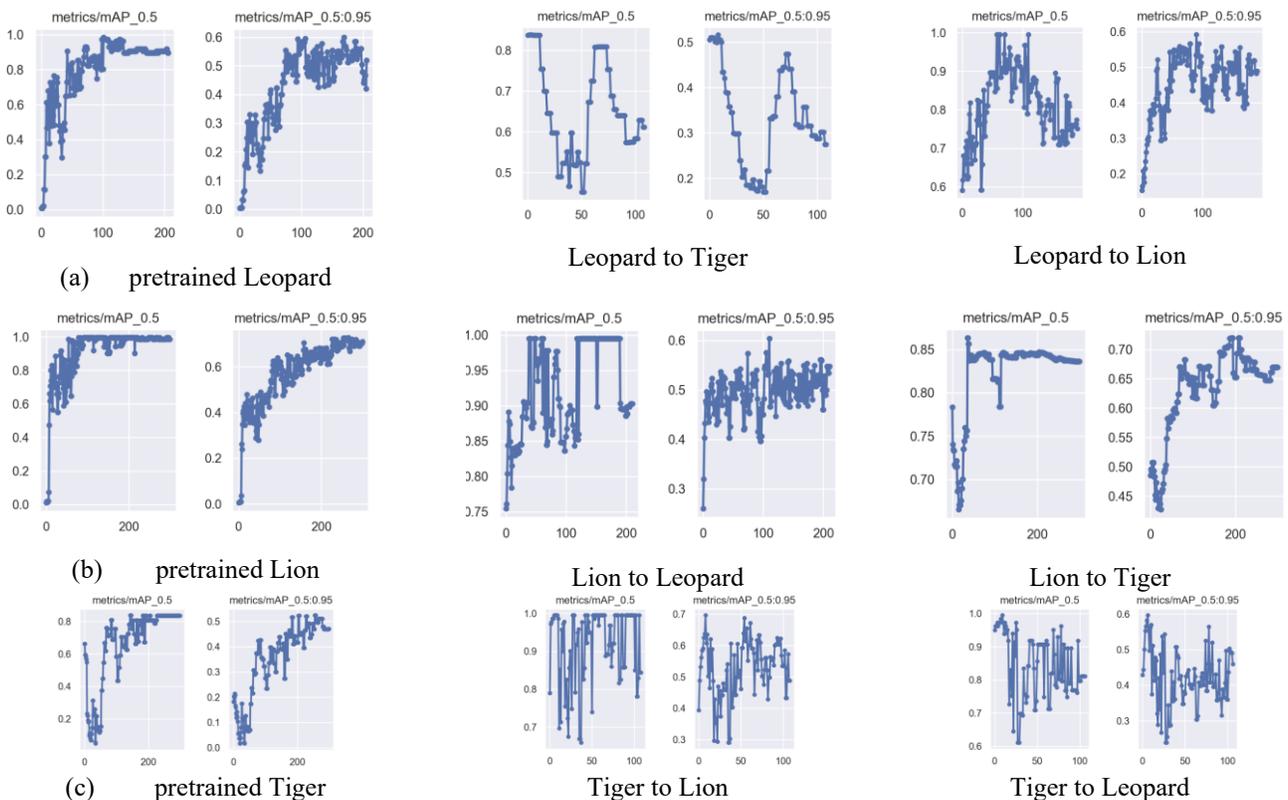
This study investigates the application of transfer learning in the classification of similar animals. By transferring knowledge from one species to another, we aim to reduce training time, improve the model's generalization capability, and achieve satisfactory classification performance even with limited data.

YOLOv5 is well-known for its benefits in terms of speed, precision, and efficiency [7]. This makes it suited for real-time monitoring in contexts such as wildlife reserves or farms [8]. The YOLOv5 design is ideal for high detection accuracy, especially in cases when animals are physically similar, such as lions, leopards, and tigers. YOLOv5's versatility across model sizes enables deployment on devices like cameras or drones, allowing for real-time inference with minimum processing resources. Additionally, its strong support for transfer learning makes it suitable for custom datasets, which helps to improve detection performance for specific animal types. YOLOv5 also incorporates advanced data augmentation techniques, enhancing its ability to generalize across varying environmental conditions, lighting, and poses. These attributes make it an ideal choice for animal intrusion detection tasks, where rapid and reliable detection is critical for timely intervention.

### 3. Proposed approach

In this study, the YOLOv5 model was utilised from the Ultralytics repository for object detection and classification of lions, leopards, and tigers [9]. The model was first pretrained on leopards and then fine-tuned for lions using transfer learning. The same procedure was later applied to lions and tigers. The main detection model is based on Convolutional Neural Networks (CNNs). The CNN is designed to classify and identify images as containing animals or not. The network contains convolutional layers followed by max-pooling layers for feature extraction and fully connected layers for classification. The project uses a custom dataset of animal images to train the model. This dataset consists of various animal images which represent different types of animal intrusions. OpenCV is used for image pre-processing tasks like resizing, image normalization, and data augmentation. This ensures the input data is standardized before feeding it into the CNN. Techniques like image thresholding and edge detection are potentially used to identify the regions where animals might be present in the captured frames. The algorithm is structured to work with camera feed input. Images are processed and analysed to determine if an animal is detected in the frame. If an animal is detected, the system is designed to trigger an alert mechanism, which can notify authorities or trigger automated actions like turning on lights or sirens to scare the animal away.

### 4. Experiment setup and results



**Figure 1.** Comparison of mAP performance (mAP@0.5 and mAP@0.5:0.95) across different transfer learning scenarios using YOLOv5 model

There are several datasets, such as The Oxford-IIIT Pet Dataset, The iNaturalist Dataset and the COCO Dataset (Common Objects in Context), used widely in animal detection and classification tasks [10]. These datasets typically contain annotated images of various species of animals, often captured in different environments and conditions.

In this study, we target at COCO Dataset [11]. Although primarily focused on objects commonly found in everyday scenes, the COCO dataset includes several animal classes and is widely used for benchmarking object detection models. Its diversity in terms of object classes, background variation, and complexity makes it a valuable resource for training and evaluating models in challenging environments.

#### 4.1. Leopard models

The pretrained leopard model achieves a peak mAP@0.5 of 0.97, showing that the model is highly effective at detecting and classifying leopards. The mAP@0.5:0.95 value of 0.57 reflects a decent performance across varying IoU thresholds, indicating the model's ability to maintain accuracy even with tighter bounding box requirements. When the leopard-pretrained model is transferred to detect tigers, the performance significantly drops. The peak mAP@0.5 reaches 0.80, while the mAP@0.5:0.95 drops further to 0.42. This indicates that despite some success in detecting tigers, the model struggles with localization precision, likely due to the visual differences between leopards and tigers (e.g., differences in patterns and body structure). The transfer from leopard to lion performs considerably better than the leopard-to-tiger transfer. The peak mAP@0.5 reaches 0.91, and the mAP@0.5:0.95 hits 0.64, suggesting that the model adapts more effectively to lions. This could be attributed to the similar body shapes and sizes of leopards and lions, making the features learned from leopards more transferable to lions.

#### 4.2. Lion models

The lion-pretrained model achieves a peak mAP@0.5 of 0.98, indicating high performance in detecting and classifying lions. The mAP@0.5:0.95 value of 0.58 indicates a robust ability to handle varying IoU thresholds, showing that the model generalizes well even in more challenging scenarios with tighter localization requirements. Transferring the lion-pretrained model to detect leopards yields excellent results, with a peak mAP@0.5 of 1.0 and a mAP@0.5:0.95 of 0.62. This strong performance suggests that lions and leopards share many visual similarities that allow for effective transfer learning. The almost perfect mAP@0.5 suggests the model adapts extremely well to leopards, while the mAP@0.5:0.95 also remains high, though slightly lower, due to the increased localization challenge. The lion-to-tiger transfer results in a peak mAP@0.5 of 0.86, which, while still solid, indicates a more challenging adaptation compared to the lion-to-leopard transfer. The mAP@0.5:0.95 value of 0.49 also shows that detecting tigers is harder for the lion-pretrained model, likely due to differences in appearance, especially in fur patterns and colouring.

#### 4.3. Tiger models

The tiger-pretrained model performs comparably to the other pretrained models, with a peak mAP@0.5 of 0.98 and a mAP@0.5:0.95 of 0.55. These results demonstrate that the model effectively detects tigers and can generalize reasonably well across varying IoU thresholds, though with some decrease in performance as the IoU threshold increases. Transferring the tiger-pre-trained model to detect lions results in a strong performance, with a peak mAP@0.5 of 0.95. The mAP@0.5:0.95, however, reaches 0.50, reflecting some struggle in tighter bounding box localization. This suggests that while the model successfully transfers to lions, the differences between tigers and lions (e.g., pattern, facial structure) affect performance at higher IoU thresholds. The tiger-to-leopard transfer yields a peak mAP@0.5 of 0.93, indicating that the model adapts well to leopard detection. However, the mAP@0.5:0.95 of 0.48 shows a similar pattern to tiger-to-lion transfer, where the model struggles more with precise localization, likely due to the differences in body shape and texture between tigers and leopards.

### 5. Conclusions and future works

Across all species, the pre-trained models already exhibit strong performance, achieving high mAP@0.5 values, indicating that each model is highly effective in detecting the species it was trained on. However, the transfer learning results vary depending on the similarity between the source and target species. Leopard-to-Lion and Lion-to-Leopard transfers show the best performance, as these species share similar physical characteristics, allowing for effective knowledge transfer. Leopard-to-Tiger and Lion-to-Tiger transfers perform moderately, indicating more difficulty in transferring between these species, likely due to their visual differences. Tiger-to-Lion and Tiger-to-Leopard transfers also demonstrate good performance but show reduced precision in higher IoU thresholds, highlighting the challenge of precise localization when transferring between species with distinct visual features. Through this work, Transfer Learning has been further verified in Animal Species Recognition tasks. It also speeds up

the training of Animal detection tasks and proves that transfer learning can be generalized in this field through experiments in Tiger, Lion, and Leopard datasets.

For future works, one of the most essential future directions is to expand the dataset to include more species, as well as increase the diversity of images per species. By incorporating more varied lighting conditions, angles, and natural environments, we can enhance the robustness of the model and improve generalization to unseen data. Additionally, adding more rare species or animals with subtle visual differences will test the model's adaptability further and provide insights into transfer learning's effectiveness across diverse species.

## References

- [1] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: <https://doi.org/10.1109/tpami.2016.2577031>.
- [2] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004, doi: <https://doi.org/10.1023/b:visi.0000029664.99615.94>.
- [3] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: <https://doi.org/10.1109/tkde.2009.191>.
- [4] A. Esteva et al., "Dermatologist-level Classification of Skin Cancer with Deep Neural Networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Jan. 2017, doi: <https://doi.org/10.1038/nature21056>.
- [5] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," *arXiv.org*, 2014. <https://arxiv.org/abs/1411.1792>
- [6] N. M. Haddad et al., "Habitat fragmentation and its lasting impact on Earth's ecosystems," *Science Advances*, vol. 1, no. 2, p. e1500052, Mar. 2015, doi: <https://doi.org/10.1126/sciadv.1500052>.
- [7] G. Jocher, "ultralytics/yolov5," *GitHub*, Aug. 21, 2020. <https://github.com/ultralytics/yolov5>
- [8] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv.org*, 2018. <https://arxiv.org/abs/1804.02767>
- [9] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv*, Apr. 2020, Available: <https://arxiv.org/abs/2004.10934>
- [10] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, Apr. 2015, doi: <https://doi.org/10.1007/s11263-015-0816-y>.
- [11] T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," *Computer Vision – ECCV 2014*, vol. 8693, pp. 740–755, 2014, doi: [https://doi.org/10.1007/978-3-319-10602-1\\_48](https://doi.org/10.1007/978-3-319-10602-1_48).