Wildlife Detection with Drones

Subjects: Computer Science, Interdisciplinary Applications Contributor: Chao Mou, Tengfei Liu, Chengcheng Zhu, Xiaohui Cui

With the impacts of global climate change and habitat loss, wild animals are facing unprecedented threats to their survival. For instance, rising temperatures, increased extreme weather events, and rising sea levels all pose threats to the survival of wildlife. The global urbanization and agricultural expansion have led to large-scale habitat destruction and loss. This has deprived many wild animals of places to live and breed. Moreover, the decline of biodiversity has also exacerbated human concerns about environmental sustainability. As a result, the issue of wildlife conservation has received extensive international attention. In this context, governments and international organizations have taken various initiatives in an attempt to curb the decline of wildlife populations and to improve the efficiency of conservation through the use of scientific and technological means. Drones are widely used for wildlife monitoring. Deep learning algorithms are key to the success of monitoring wildlife with drones, although they face the problem of detecting small targets.

Keywords: drones ; wildlife detection ; gated channel attention mechanism ; image dataset ; YOLOv7 ; small target detection

1. Introduction

In recent years, aerial imagery has been extensively utilized in wildlife conservation due to the fast-paced evolution of drones and other aerial platforms ^[1]. Moreover, Unmanned Aerial Vehicles (UAVs) offer several advantages in aerial imagery due to their high altitude and ability to carry high-resolution camera payloads. Aerial imagery can, therefore, provide humans with valuable information in complex and harsh natural environments in a relatively short period of time ^[2]. For instance, high-resolution aerial photographs can accurately identify wildlife categories and population numbers, facilitating wildlife population counts and assessments of their endangerment levels [3]. Similarly, aerial photographs with broader coverage can detect wildlife in remote or otherwise inaccessible areas ^[4]. The wide field of view is advantageous as it enables the aerial images to observe a large area, covering several square kilometers, while simultaneously detecting individuals and groups of animals that are otherwise difficult to access [5]. Various methods have been explored for detecting wildlife on aerial images, including target detection algorithms, semantic segmentation algorithms, and deep learning methods [6][7][8][9][10]. For instance, Barbedo et al. [7] used convolutional neural networks to monitor herds of cattle with drones. Brown et al. [8] employed high-resolution UAV images to establish an accurate ground truth. They utilized a target detection-based method using YOLOv5 to achieve animal localization and counting on Australian farms. Similarly, Padubidri et al. [10] counted sea lions and African elephants in aerial images by utilizing a UNet model based on semantic segmentation. Furthermore, wildlife detection using deep learning techniques and aerial imagery can rapidly locate wildlife, aiding in the planning of effective protection and search activities [11]. Additionally, it can determine the living conditions of wildlife in dangerous and complex natural habitats such as swamps and forests [12]. In a word, deep learning models, such as YOLO, have emerged as a significant approach in UAV aerial imagery-based applications for wildlife detection.

The success of deep learning in wildlife detection with drones relies on having a large amount of real-world data available for training algorithms ^[13]. For example, training on large datasets like ImageNet ^[14] and MSCOCO ^[15] was then followed by transfer learning on actual data. Unfortunately, the extremely high cost of acquiring high-quality aerial images makes obtaining real transfer learning data challenging, not to mention constructing large-scale aerial wildlife datasets. That is to say, due to the lack of large-scale aerial wildlife image datasets, most current methods involve adapting object detection algorithms developed for natural scene images to aerial images, which is not suitable for wildlife detection ^[16]. These challenges contribute to the current shortcomings of drone-based wildlife detection algorithms, including low accuracy, weak robustness, and unsatisfactory practical application outcomes ^{[17][18]}.

In addition, the data of wildlife captured by drones have unique characteristics. As can be seen from **Figure 1**, the individual animals in aerial images are small and the proportion varies greatly. This arises not only from differences in drone flight altitudes but also relates to inherent variations in object sizes within the same category. Therefore, it is

necessary to overcome the small target problem when using drones to detect wildlife ^[19]. Currently, researchers have invested substantial efforts in enhancing the accuracy of algorithms for detecting these small targets, but few of them focus on wildlife aerial imagery applications ^[20]. For instance, Chen et al. ^[21] proposed an approach that combines contextual models and small region proposal generators to enhance the R-CNN algorithm, thereby improving the detection of small objects in images. Zhou et al. ^[22] introduced a scale-transferable detection network that maintains scale consistency between different detection scales through embedded super-resolution layers, enhancing the detection to a certain extent, challenges remain, including limited real-time performance, high computational demands, information redundancy, and complex network structures.



Figure 1. Example of aerial images of zebra migration at different flight heights.

On the one hand, to address real-time wildlife detection, researchers often employ one-stage YOLO methods ^[23]. For example. Zhong et al. [24] used the YOLO method to monitor marine animal populations, demonstrating that YOLO can rapidly and accurately identify marine animals within coral reefs. Roy et al. [25] proposed a high-performance YOLO-based automatic detection model for real-time detection of endangered wildlife. Although advanced versions like YOLOv7 [26] have started addressing multi-scale detection issues, they still struggle with the highly dynamic scenes encountered in drone aerial imagery [27]. On the other hand, researchers have also turned their attention to attention mechanisms to address the small target detection problem. These mechanisms guide the model to focus on critical feature regions of the target, leading to more accurate localization of small objects and reduced computational resource consumption ^[28]. For instance, Wang et al. [29] introduced a pre-detection concept based on attention mechanisms into the model, which restricts the detection area in images based on the characteristics of small object detection. This reduces redundant information and improves the efficiency of small target detection. Zuo et al. [30] proposed an attention-fused feature pyramid network to address the issue of potential target loss in deep network structures. Their adaptive attention fusion module enhances the spatial positions and semantic information features of infrared small targets in shallow and deep layers, enhancing the network's ability to represent target features. Zhu et al. [31] introduced a lightweight small target detection network with embedded attention mechanisms. This enhances the network's ability to focus on regions of interest in images while maintaining a lightweight network structure, thereby improving the detection of small targets in complex backgrounds. These studies indicate that incorporating attention mechanisms into models is beneficial for effectively fusing features from different levels and enhancing the detection efficiency and capabilities of small targets. Notably, channel attention mechanisms hold advantages in small target detection [32][33]. Channel attention mechanisms better consider information from different channels, focusing on various dimensions of small targets and effectively capturing their features [34]. Moreover, since channel attention mechanisms are unaffected by spatial dimensions, they might be more stable when handling targets with different spatial distributions ^[35]. This can enhance robustness when dealing with targets of varying sizes, positions, and orientations.

2. Diverse Datasets for Wildlife Detection

Currently, there are numerous publicly available animal datasets based on ground images as shown in **Table 1**. As seen from **Table 1**, Song et al. ^[36] introduced the Animal10N dataset, containing 55,000 images from five pairs of animals: (cat, lynx), (tiger, cheetah), (wolf, coyote), (gorilla, chimpanzee), and (hamster, guinea pig). The training set comprises 50,000 images, whereas the test set has 5000 images. Ng et al. ^[37] presented the Animal Kingdom dataset, a vast and diverse collection offering annotated tasks to comprehensively understand natural animal behavior. It includes 50 h of annotated videos for locating relevant animal behavior segments in lengthy videos, 30 K video sequences for fine-grained multi-label action recognition tasks, and 33 K frames for pose estimation tasks across 850 species of six major animal phyla. Cao et al. ^[38] introduced the Animal-Pose dataset for animal pose estimation, covering five categories: dogs, cats, cows, horses, and sheep, with over 6000 instances from 4000+ images. Additionally, it provides bounding box annotations for seven animal categories: otter, lynx, rhinoceros, hippopotamus, gorilla, bear, and antelope. AnimalWeb ^[39] is a large-scale, hierarchically annotated animal face dataset, encompassing 22.4 K faces from 350 species and 21 animal families. These facial images were captured in the wild under field conditions and annotated with nine key facial landmarks consistently across the dataset. However, these existing datasets may not sufficiently support the demands of comprehensive aerial wildlife monitoring. For instance, the Animal10 dataset focuses on animal detection rather than wildlife detection in natural

environments. The Animal Kingdom and Animal-Pose datasets were not specifically designed for wildlife detection in drone aerial images.

Dataset	Animal10	Animal-Pose	AnimalWeb	Animal Kingdom	WAID (Ours)
Year	2019	2019	2020	2022	2023
multicategory	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Wilderness	×	\checkmark	\checkmark	\checkmark	\checkmark
Multiscale	×	\checkmark	×	\checkmark	\checkmark
ValidationSet	×	×	×	\checkmark	\checkmark
AerialView	×	×	×	×	\checkmark

Table 1. A summary of the animal image datasets currently available.

In addition, most publicly available aerial image datasets are concentrated in urban areas and primarily used for tasks like building detection, vehicle detection or segmentation, and road planning ^{[40][41][42]}. However, to the best of the researcher's knowledge, there are only a few public datasets available for animal detection in aerial images, such as the NOAA Arctic Seal dataset ^[43] and the Aerial Elephant Dataset (AED) ^[44]. The NOAA Arctic Seal dataset comprises approximately 80,000 color and infrared (thermal) images, collected during flights conducted by the NOAA Alaska Fisheries Science Center in Alaska in 2019. The images are annotated with around 28,000 ice seal bounding boxes (14,000 on color images and 14,000 on thermal images). The AED is a challenging dataset containing 2074 images, including a total of 15,581 African forest elephants in their natural habitat. The imaging was conducted with consistent methods across a range of background types, resolutions, and times of day. However, these datasets are annotated specifically for certain species. In addition to this, there are a number of small-scale aerial wildlife datasets (e.g., elephants or camels) available on data collection platforms (e.g., rebowflow). In summary, there are three limitations to the current aerial wildlife detection datasets. First, the current aerial wildlife datasets are small in size and not species-rich. Second, these datasets often suffer from limited image quantity and inconsistent annotation quality. Third, there are no open-source, large-scale, high-quality wildlife aerial photography datasets yet. In summary, the current aerial wildlife datasets, these limitations have restricted their applicability in diverse wildlife monitoring scenarios.

3. Object Detection Methods

3.1. Single-Stage Object Detection Method

Single-stage object detection methods refer to those that require only one feature extraction process to achieve object detection. Compared to multi-stage detection methods, single-stage methods are faster in terms of speed, although they might have slightly lower accuracy. Some typical single-stage object detection methods include the YOLO series, SSD, and RefineDet. The YOLO model was initially introduced by Redmon et al. [45] in 2015. Its core idea is to divide the image into S*S grids, and each grid is responsible for predicting and generating multiple bounding boxes if the center point of the object is within that grid. However, during training, each bounding box is allowed to predict only one object, limiting its ability to detect multiple objects and thus affecting detection accuracy. Nonetheless, it significantly improves detection speed. Shortly after the publication of the YOLO model, the SSD model was also introduced [46]. SSD uses VGG16 [47] as its backbone network and does not generate candidate boxes. Instead, it allows feature extraction at multiple scales. The SSD algorithm trains the model using a weighted combination of data augmentation, localization loss, and confidence loss, resulting in fast speeds. However, due to its challenging training process, its accuracy is relatively lower. In 2018, Ref. ^[48] proposed the RefineDet detection algorithm. This algorithm comprises the ARM module, TCB module, and ODM module. The ARM module focuses on refining anchor boxes, reducing the proportion of negative samples, and adjusting the size and position of anchor boxes to provide better results for regression and classification. The TCB module transforms the output of the ARM into the input of ODM, achieving the fusion of high-level semantic information and lowlevel feature information. RefineDet uses two-stage regression to improve object detection accuracy and achieve end-toend multi-task training.

3.2. Two-Stage Object Detection Method

Compared to single-stage detectors, two-stage object detection algorithms divide the detection process into two main steps: region proposal generation and object classification. The advantage of this approach lies in its ability to select multiple candidate boxes, effectively extracting feature information from targets, resulting in high-precision detection and

accurate localization. However, due to the involvement of two independent steps, this can lead to slower detection speeds and more complex models. Girshick et al. ^[49] introduced the classic R-CNN algorithm, which employed a "Region Proposal + CNN" approach to feature extraction, opening the door for applying deep learning to object detection and laying the foundation for subsequent algorithms. Subsequently, He et al. ^[50] introduced the SPP-Net algorithm to address the time-consuming feature extraction process in R-CNN. To further improve upon the drawbacks of R-CNN and SPP-Net, Girshick ^[51] proposed the Fast R-CNN algorithm. Fast R-CNN combined the advantages of R-CNN and SPP-Net, but still did not achieve end-to-end object detection. In 2015, Ren et al. ^[52] introduced the Faster R-CNN algorithm, which was the first deep learning-based detection algorithm that approached real-time performance. Faster R-CNN introduced Region Proposal Networks (RPNs) to replace traditional Selective Search algorithms for generating proposals. Although Faster R-CNN surpassed the speed limitations of Fast R-CNN, there still remained issues of computational redundancy in the subsequent detection. Furthermore, Faster R-CNN's performance in detecting small objects was subpar. Various improvements were subsequently proposed, such as R-FCN ^[53] and Light-Head R-CNN ^[54], which further enhanced the performance of Faster R-CNN.

4. Gated Channel Attention Mechanism

Incorporating attention mechanisms into neural networks offers various methods, and taking convolutional neural networks (CNNs) as an example, attention can be introduced in the spatial dimension, as well as in the channel dimension. For instance, in the Inception ^[55] network, the multi-scale approach assigns different weights to parallel convolutional layers, offering an attention-like mechanism. Attention can also be introduced in the channel dimension ^[56], and it is possible to combine both spatial and channel attention mechanisms ^[57]. The channel attention mechanism focuses on the differences in importance among channels within a convolutional layer, adjusting the weights of channels to enhance feature extraction. On the other hand, the spatial attention mechanism builds upon the channel attention concept, suggesting that the importance of each pixel in different channels varies. By adjusting the weights of all pixels across different channels, the network's feature extraction capabilities are enhanced.

When applied to small object detection in aerial images, the channel attention mechanism is considered more suitable and has seen substantial research in enhancing small object detection while also addressing model complexity. For example, Wang et al. ^[58] addressed the trade-off between model detection performance and complexity by proposing an efficient channel attention module (ECA), which introduces a small number of parameters but yields significant performance improvements. Tong et al. ^[59] introduced a channel attention-based DenseNet network that rapidly and accurately captures key features from images, thus improving the classification of remote sensing image scenes. To tackle the issue of low recognition rates and high miss rates in current object detection tasks, Yang et al. ^[60] proposed an improved YOLOv3 algorithm incorporating a gated channel attention mechanism (GCAM) and an adaptive upsampling module. Results showed that the improved approach adapts to multi-scale object detection tasks in complex scenes and reduces the omission rate of small objects. Hence, in this research, a channel attention mechanism to enhance the performance of YOLO with drones is attempted use.

5. Drone-Based Wildlife Detection

In the field of wildlife detection using drones, UAVs have made wildlife conservation technology more accessible. The benefits of aerial imagery from drones have contributed to wildlife conservation.

Drones have automated the process of capturing high-altitude aerial photos of wildlife habitats, greatly simplifying the process of obtaining images and monitoring wildlife using automated techniques such as image recognition. This technology eliminates the necessity for capturing images manually and enables more efficient monitoring of wildlife populations. For example, Hodgson et al. ^[61] applied UAV technology to wildlife monitoring in tropical and polar environments and demonstrated through their study that UAV counts of nesting bird colonies were an order of magnitude more accurate than traditional ground counts. Sirmacek et al. ^[62] proposed a computer vision method for quantifying animals in aerial images, resulting in successful detection and counting. The approach holds promise for effective wildlife conservation management. In a study comparing the efficacy of object- and pixel-based classification methods, Ruwaimana et al. ^[63] established that UAV images were more informative than satellite images in mapping mangrove forests in Malaysian wetlands. This opens the possibility for automating the detection and analysis of wildlife in aerial images using computer vision techniques, hence rendering wildlife conservation studies based on these images more efficient and feasible. Raoult et al. ^[64] employed drones in examining marine animals, which enabled the study of numerous aspects of ecology, behavior, health, and movement patterns of these animals. Overall, UAVs and aerial imagery have significant research value and potential for development in wildlife conservation.

Moreover, aerial imagery obtained through the use of drones is useful in various wildlife conservation applications due to its high level of accuracy, vast coverage, and broad field of view. This type of imagery, particularly those with 4K resolution or higher, can accurately identify various animal details, such as their category and quantity ^[65]. For instance, the distinctive stripes of zebras can easily be recognized ^[66], which is essential in counting the number of animals and assessing their level of endangerment. Aerial photos that cover a broad territory can be useful to survey wildlife in remote or hard-to-reach locations, aiding in complete assessments of animal distribution and habitat utilization ^[67]. For the advantage of comprehensive coverage, UAVs can survey areas encompassing vast wetlands or dense forests to detect elusive wildlife populations and monitor their ranges. This helps develop effective conservation plans ^[68].

However, current algorithms for wildlife detection using UAVs still encounter various challenges. Firstly, the accuracy of detecting small targets is inadequate. Existing algorithms fail to satisfy the requirements for accurate detection of small targets. As UAVs capture images from far away, small targets are often only a few tens of pixels. For instance, an enhanced YOLOv4 model was proposed by Tan et al. ^[69] for detecting UAV aerial images, which achieved only 45% of the mean average precision. Wang et al. [12] discovered that identifying small animals within obstructed jungle environments during UAV wildlife investigations is challenging. Furthermore, meeting practical use standards requires enhanced real-time algorithm performance, which current techniques struggle to achieve the necessary response speed for flight. For instance, Benjdira et al. [70] utilized the Faster R-CNN method in detecting vehicles in aerial images. The experimental results revealed that the detection speed hit 1.39 s. However, this could not satisfy the real-time requirements. Fortunately, YOLO, a one-stage detection algorithm, exhibits outstanding real-time performance, making it a crucial choice as the primary model for this study. Additionally, the publicly accessible training data remain limited, which inevitably leads to the model's overfitting issue and weak generalization ability in practical scenarios. These limitations suggest that there remains a disparity between the existing algorithms and the necessity for ongoing enhancement in order to attain precise and effective identification of wildlife in aerial images. Consequently, it has been spurred to develop a real-time detection (e.g., YOLO) technique for detecting small-target UAV wildlife and to divulge a superior, open-source training dataset to advance the field of academia.

References

- Descamps, S.; Béchet, A.; Descombes, X.; Arnaud, A.; Zerubia, J. An automatic counter for aerial images of aggregations of large birds. Bird Study 2011, 58, 302–308.
- Ševo, I.; Avramović, A. Convolutional neural network based automatic object detection on aerial images. IEEE Geosci. Remote Sens. Lett. 2016, 13, 740–744.
- 3. Chabot, D.; Francis, C.M. Computer-automated bird detection and counts in high-resolution aerial images: A review. J. Field Ornithol. 2016, 87, 343–359.
- Li, X.; Xing, L. Use of unmanned aerial vehicles for livestock monitoring based on streaming K-means clustering. IFAC-PapersOnLine 2019, 52, 324–329.
- Sundaram, D.M.; Loganathan, A. FSSCaps-DetCountNet: Fuzzy soft sets and CapsNet-based detection and counting network for monitoring animals from aerial images. J. Appl. Remote Sens. 2020, 14, 026521.
- Ward, S.; Hensler, J.; Alsalam, B.; Gonzalez, L.F. Autonomous UAVs wildlife detection using thermal imaging, predictive navigation and computer vision. In Proceedings of the 2016 IEEE Aerospace Conference, Big Sky, MT, USA, 5–12 March 2016; pp. 1–8.
- 7. Barbedo, J.G.A.; Koenigkan, L.V.; Santos, T.T.; Santos, P.M. A study on the detection of cattle in UAV images using deep learning. Sensors 2019, 19, 5436.
- 8. Brown, J.; Qiao, Y.; Clark, C.; Lomax, S.; Rafique, K.; Sukkarieh, S. Automated aerial animal detection when spatial resolution conditions are varied. Comput. Electron. Agric. 2022, 193, 106689.
- Hong, S.J.; Han, Y.; Kim, S.Y.; Lee, A.Y.; Kim, G. Application of deep-learning methods to bird detection using unmanned aerial vehicle imagery. Sensors 2019, 19, 1651.
- 10. Padubidri, C.; Kamilaris, A.; Karatsiolis, S.; Kamminga, J. Counting sea lions and elephants from aerial photography using deep learning with density maps. Anim. Biotelem. 2021, 9, 27.
- 11. Linchant, J.; Lisein, J.; Semeki, J.; Lejeune, P.; Vermeulen, C. Are unmanned aircraft systems (UASs) the future of wildlife monitoring? A review of accomplishments and challenges. Mammal Rev. 2015, 45, 239–252.
- 12. Wang, D.; Shao, Q.; Yue, H. Surveying wild animals from satellites, manned aircraft and unmanned aerial systems (UASs): A review. Remote Sens. 2019, 11, 1308.

- 13. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. J. Big Data 2016, 3, 9.
- Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
- Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft Coco: Common objects in context. In Proceedings of the Computer Vision—ECCV 2014, 13th European Conference, Zurich, Switzerland, 6–12 September 2014; Springer: Cham, Switzerland, 2014; pp. 740–755.
- Zheng, X.; Kellenberger, B.; Gong, R.; Hajnsek, I.; Tuia, D. Self-supervised pretraining and controlled augmentation improve rare wildlife recognition in UAV images. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 732–741.
- Okafor, E.; Smit, R.; Schomaker, L.; Wiering, M. Operational data augmentation in classifying single aerial images of animals. In Proceedings of the 2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA), Gdynia, Poland, 3–5 July 2017; pp. 354–360.
- Kellenberger, B.; Marcos, D.; Tuia, D. Best practices to train deep models on imbalanced datasets—A case study on animal detection in aerial imagery. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Dublin, Ireland, 10–14 September 2018; Springer: Cham, Switzerland, 2018; pp. 630–634.
- 19. Wang, Y.; Han, D.; Wang, L.; Guo, Y.; Du, H. Contextualized Small Target Detection Network for Small Target Goat Face Detection. Animals 2023, 13, 2365.
- 20. Gao, C.; Meng, D.; Yang, Y.; Wang, Y.; Zhou, X.; Hauptmann, A.G. Infrared patch-image model for small target detection in a single image. IEEE Trans. Image Process. 2013, 22, 4996–5009.
- Chen, C.; Liu, M.Y.; Tuzel, O.; Xiao, J. R-CNN for small object detection. In Proceedings of the Computer Vision— ACCV 2016, 13th Asian Conference on Computer Vision, Taipei, Taiwan, 20–24 November 2016; Springer: Cham, Switzerland, 2017; pp. 214–230.
- 22. Zhou, P.; Ni, B.; Geng, C.; Hu, J.; Xu, Y. Scale-transferrable object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 528–537.
- 23. Pham, M.T.; Courtrai, L.; Friguet, C.; Lefèvre, S.; Baussard, A. YOLO-Fine: One-stage detector of small objects under various backgrounds in remote sensing images. Remote Sens. 2020, 12, 2501.
- 24. Zhong, J.; Li, M.; Qin, J.; Cui, Y.; Yang, K.; Zhang, H. Real-time marine animal detection using YOLO-based deep learning networks in the coral reef ecosystem. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2022, 46, 301–306.
- 25. Roy, A.M.; Bhaduri, J.; Kumar, T.; Raj, K. WilDect-YOLO: An efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection. Ecol. Inform. 2023, 75, 101919.
- Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 7464–7475.
- He, Y.; Su, B.; Yan, J.; Tang, J.; Liu, C. Research on underwater object detection of improved YOLOv7 model based on attention mechanism: The underwater detection module YOLOv7-C. In Proceedings of the 2022 4th International Conference on Robotics, Intelligent Control and Artificial Intelligence, Dongguan, China, 16–18 December 2022; pp. 302–307.
- 28. Pashler, H.; Johnston, J.C.; Ruthruff, E. Attention and performance. Annu. Rev. Psychol. 2001, 52, 629-651.
- 29. Wang, Y.; Zhang, T.; Wang, G. Small-target predetection with an attention mechanism. Opt. Eng. 2002, 41, 872–885.
- 30. Zuo, Z.; Tong, X.; Wei, J.; Su, S.; Wu, P.; Guo, R.; Sun, B. AFFPN: Attention fusion feature pyramid network for small infrared target detection. Remote Sens. 2022, 14, 3412.
- Zhu, W.; Wang, L.; Jin, Z.; He, D. Lightweight small object detection network with attention mechanism. Opt. Precis. Eng. 2022, 30, 998–1010.
- Kim, M.; Jeong, J.; Kim, S. ECAP-YOLO: Efficient channel attention pyramid YOLO for small object detection in aerial image. Remote Sens. 2021, 13, 4851.
- Dai, Y.; Gieseke, F.; Oehmcke, S.; Wu, Y.; Barnard, K. Attentional feature fusion. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2021; pp. 3560–3569.
- 34. Guo, C.; Fan, B.; Zhang, Q.; Xiang, S.; Pan, C. AugFPN: Improving multi-scale feature learning for object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 12595–12604.

- 35. Wang, Z.; Wang, B.; Xu, N. SAR ship detection in complex background based on multi-feature fusion and non-local channel attention mechanism. Int. J. Remote Sens. 2021, 42, 7519–7550.
- 36. Song, H.; Kim, M.; Lee, J.G. Selfie: Refurbishing unclean samples for robust deep learning. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 5907–5915.
- 37. Ng, X.L.; Ong, K.E.; Zheng, Q.; Ni, Y.; Yeo, S.Y.; Liu, J. Animal Kingdom: A large and diverse dataset for animal behavior understanding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–24 June 2022; pp. 19023–19034.
- Cao, J.; Tang, H.; Fang, H.S.; Shen, X.; Lu, C.; Tai, Y.W. Cross-domain adaptation for animal pose estimation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9498–9507.
- 39. Khan, M.H.; McDonagh, J.; Khan, S.; Shahabuddin, M.; Arora, A.; Khan, F.S.; Shao, L.; Tzimiropoulos, G. Animalweb: A large-scale hierarchical dataset of annotated animal faces. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 6939–6948.
- Maggiori, E.; Tarabalka, Y.; Charpiat, G.; Alliez, P. Can semantic labeling methods generalize to any city? The inria aerial image labeling benchmark. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 3226–3229.
- 41. Chen, Q.; Wang, L.; Wu, Y.; Wu, G.; Guo, Z.; Waslander, S. Aerial imagery for roof segmentation: A large-scale dataset towards automatic mapping of buildings. arXiv 2018, arXiv:1807.09532.
- 42. Van Etten, A.; Lindenbaum, D.; Bacastow, T.M. Spacenet: A remote sensing dataset and challenge series. arXiv 2018, arXiv:1807.01232.
- Lee Son, G.; Romain, S.; Rose, C.; Moore, B.; Magrane, K.; Packer, P.; Wallace, F. Development of Electronic Monitoring (EM) Computer Vision Systems and Machine Learning Algorithms for Automated Catch Accounting in Alaska Fisheries; Alaska Fisheries Science Center; NOAA; National Marine Fisheries Service: Seattle, WA, USA, 2023.
- 44. Naude, J.; Joubert, D. The Aerial Elephant Dataset: A New Public Benchmark for Aerial Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–20 June 2019; pp. 48–55.
- Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
- 46. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the Computer Vision—ECCV 2016, 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; pp. 21–37.
- 47. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.
- Zhang, S.; Wen, L.; Bian, X.; Lei, Z.; Li, S.Z. Single-shot refinement neural network for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4203–4212.
- Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
- 50. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2015, 37, 1904–1916.
- Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7– 13 December 2015; pp. 1440–1448.
- 52. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. Adv. Neural Inf. Process. Syst. 2015, 28, 91–99.
- Dai, J.; Li, Y.; He, K.; Sun, J. R-FCN: Object detection via region-based fully convolutional networks. Adv. Neural Inf. Process. Syst. 2016, 29, 379–387.
- 54. Li, Z.; Peng, C.; Yu, G.; Zhang, X.; Deng, Y.; Sun, J. Light-head R-CNN: In defense of two-stage object detector. arXiv 2017, arXiv:1711.07264.
- 55. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June

2016; pp. 2818-2826.

- Yu, C.; Wang, J.; Peng, C.; Gao, C.; Yu, G.; Sang, N. BiSeNet: Bilateral segmentation network for real-time semantic segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 325–341.
- 57. Fu, J.; Liu, J.; Jiang, J.; Li, Y.; Bao, Y.; Lu, H. Scene segmentation with dual relation-aware attention network. IEEE Trans. Neural Netw. Learn. Syst. 2020, 32, 2547–2560.
- Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11534–11542.
- 59. Tong, W.; Chen, W.; Han, W.; Li, X.; Wang, L. Channel-attention-based DenseNet network for remote sensing image scene classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2020, 13, 4121–4132.
- 60. Yang, X.; Shi, J.; Zhang, J. Gated channel attention mechanism YOLOv3 network for small target detection. Adv. Multimed. 2022, 2022, 8703380.
- 61. Hodgson, J.C.; Baylis, S.M.; Mott, R.; Herrod, A.; Clarke, R.H. Precision wildlife monitoring using unmanned aerial vehicles. Sci. Rep. 2016, 6, 22574.
- 62. Sirmacek, B.; Wegmann, M.; Cross, A.; Hopcraft, J.; Reinartz, P.; Dech, S. Automatic population counts for improved wildlife management using aerial photography. In Proceedings of the 6th International Congress on Environmental Modelling and Software, Leipzig, Germany, 1–5 July 2012.
- 63. Ruwaimana, M.; Satyanarayana, B.; Otero, V.; Muslim, A.M.; Syafiq, A.M.; Ibrahim, S.; Raymaekers, D.; Koedam, N.; Dahdouh-Guebas, F. The advantages of using drones over space-borne imagery in the mapping of mangrove forests. PLoS ONE 2018, 13, e0200288.
- Raoult, V.; Colefax, A.P.; Allan, B.M.; Cagnazzi, D.; Castelblanco-Martínez, N.; Ierodiaconou, D.; Johnston, D.W.; Landeo-Yauri, S.; Lyons, M.; Pirotta, V.; et al. Operational protocols for the use of drones in marine animal research. Drones 2020, 4, 64.
- Mnih, V.; Hinton, G.E. Learning to detect roads in high-resolution aerial images. In Proceedings of the Computer Vision —ECCV 2010, 11th European Conference on Computer Vision, Heraklion, Crete, Greece, 5–11 September 2010; Springer: Cham, Switzerland, 2010; pp. 210–223.
- Fang, Y.; Du, S.; Abdoola, R.; Djouani, K.; Richards, C. Motion based animal detection in aerial videos. Procedia Comput. Sci. 2016, 92, 13–17.
- 67. Bennitt, E.; Bartlam-Brooks, H.L.; Hubel, T.Y.; Wilson, A.M. Terrestrial mammalian wildlife responses to Unmanned Aerial Systems approaches. Sci. Rep. 2019, 9, 2142.
- 68. Fust, P.; Loos, J. Development perspectives for the application of autonomous, unmanned aerial systems (UASs) in wildlife conservation. Biol. Conserv. 2020, 241, 108380.
- 69. Tan, L.; Lv, X.; Lian, X.; Wang, G. YOLOv4_Drone: UAV image target detection based on an improved YOLOv4 algorithm. Comput. Electr. Eng. 2021, 93, 107261.
- Benjdira, B.; Khursheed, T.; Koubaa, A.; Ammar, A.; Ouni, K. Car detection using unmanned aerial vehicles: Comparison between faster R-CNN and YOLOv3. In Proceedings of the 2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS), Muscat, Oman, 5–7 February 2019; pp. 1–6.

Retrieved from https://encyclopedia.pub/entry/history/show/113416