Data Augmentation with 3D-Rendered Models for Livestock Recognition Using Drone Footage

Kein Yamada Takahiro Shindo Taiju Watanabe Hiroshi Watanabe
Graduate School of Fundamental Science and Engineering, Waseda University

Abstract

Livestock counting or recognition is a crucial process of cattle management in every farm worldwide. Farmland owners typically use drones to capture instances or footage of livestock, especially when the livestock count is massive. Livestock recognition of drone footage requires many variations of aerial images or videos of livestock for training, which are generally insufficient in real life. Our paper demonstrates a method of data augmentation using 3D models and scenes rendered for the training dataset of livestock recognition. Experimental results in our research convey a satisfying detection result, validating the use of 3D-rendered models for data augmentation.

Keywords: livestock management, data augmentation, image recognition

1. Introduction

Livestock management is a significant yet difficult task with many objectives, holding variety of challenges for maintaining an optimal farm environment. In this paper, we will focus on the objective of counting, detecting, and tracking livestock in farms. The use of drone devices is fitting for such purpose, due to its ability of effectively collecting footage in a form of bird's eye, through the attached camera. Drones are also evaluated to have great contribution on counting, detecting, and tracking the animals that are visible on the camera [1]. However, drones cannot recognize the presence of the livestock on its own. UAVs require software or tools that has an object detection model with trained parameters, requiring a variety of drone footage in a form of bird's view, which are often insufficient. To confront this concern, we formulate a data augmentation method that utilizes Blender to merge 3D model of an animal and 3D scenes, rendered as the replacement for the training dataset of the livestock recognition [2].

2. Related Work

Data augmentation is a commonly conducted process for increasing the diversity of a specific dataset, mainly for the purpose of preventing over-fitting in object detection models. Typically, data augmentation is involves cropping and rotating of the images. A GitHub repository by Lopez, also uses Blender to automatically generate images and bounding box labels [3]. However, the repository is not meant for tracking the livestock since animals do not remain still.

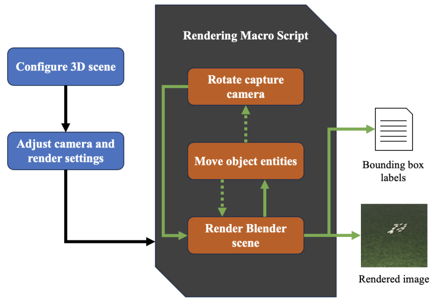


Figure 1: Flowchart showcasing the procedures and steps taken for the process of Blender-based data augmentation.

3. Proposed Method

First, we choose a sheep 3D model (designed by Creazilla, https://creazilla.com) to work with. Next, we construct a Blender environment for rendering the images, such as configuring the 3D scene, rendering, and the camera. Once the Blender environment and the sheep are properly set up, we use a rendering macro script to automatically create a 3D sheep model-based dataset and the corresponding bounding box labels, similar to what Lopez's repository does. The difference between our method and Lopez's method is the addition of object movements. This helps with providing variation of sheep moving and also for task such as object tracking, as shown on the right of Fig. 2. The movement is added in between the rendering process and camera rotation process, as shown in Fig. 1. This addition provides object movements for every single angle captured in Blender. These steps of the rendering macro script are repeated until we get 3000 images of bird's eye view sheep, utilized for the training dataset, shown on the left of Fig. 2.

4. Experiment

4.1. Evaluation

For evaluation, we first train the dataset on the object detection model, YOLOv5 [4]. We use the dataset uploaded

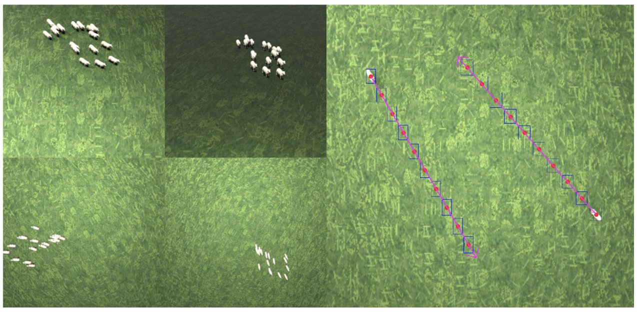


Figure 2: Left side shows the images achieved from the blender rendering macro. Right side shows the application of the method for object tracking.

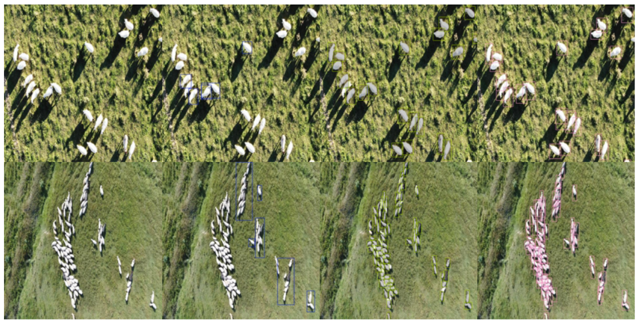
Table 1: Sheep detection performance comparison of different detection models [%]

	YOLOv5	RIIS	Ours
FNR ↓	99.6	**8.6**	13.4
TPR ↑	0.4	**91.4**	86.6
Accuracy ↑	0.4	**91.4**	84.2

on Roboflow (RIIS) for validation, containing 350 images [5] [6]. We proceed with fine-tuning YOLOv5m for model weight. For detection, we choose 10 images from the RIIS dataset, which are passed through detection models with a confidence threshold of 0.5 to calculate the false negative rate (FNR), true positive rate (TPR), and the accuracy of the detection results. This is shown in Fig. 3 where YOLOv5, RIIS, and Ours refers to results from YOLOv5m weight, results produced from Roboflow detection model trained using the RIIS dataset, and results achieved from weight trained on proposed dataset, respectively.

4.2. Result

The results are obtained as shown above in Table 1. Looking at the results from YOLOv5, it can be seen that the model's detection has an extremely high FNR. This includes object detection such as correctly placed bounding box but incorrect class label or multiple objects seen as one object. The TPR and accuracy are the same for YOLOv5, given as 0.4 %, since there were no false positives. Our proposed method has a significantly reduced FNR of 11.3 % compared to the result of YOLOv5. Great improvements are also seen in the TPR and accuracy, with an increase of 86.2 % and 83.8 %, respectively. Unfortunately, our results are lower than RIIS by about 4.8 % for TPR and 7.2 % for accuracy. However, these results are meaningful as a proof of improvement in sheep detection.



(a) **(b)** **(c)** **(d)**

Figure 3: The livestock detection results are shown for clarity. (a) original image; (b) YOLOv5; (c) RIIS; (d) Ours.

5. Conclusion

We introduced a data augmentation method of generating images from 3D models and scenes by utilizing blender, which created drone view images of sheep. Results have shown that our proposed method is capable for counting and detecting sheep compared to the initial state of the YOLOv5 model. Furthermore, object tracking can be done, due to the addition of object movement and the macro that automatically gives us the center of the object and the bounding box coordinates. In our future research, we intend to replace the 3D object models to animals other than sheep for more accuracy measurements. We also seek to make minor improvements to minimize the FP and FN count, ultimately for detecting larger herd of sheep.

Acknowledgment

These research results were obtained from the commissioned research (JPJ012368C05101) by National Institute of Information and Communications Technology (NICT), Japan.

References

[1] M. A. Alanezi, A. F. Salami, Y. A. Sha'aban, *et al.*, "Ruber: Recoverable uav-based energy-efficient reconfigurable routing scheme for smart wireless livestock sensor network," *Frontiers in Energy Research*, vol. 10, 2022, DOI: 10.3389/fenrg.2022.975614.

[2] B. O. Community, *Blender - a 3d modelling and rendering package*, Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. [Online]. Available: http://www.blender.org.

[3] F.A. Lopez, *federicoarenasl/Data-Generation-with-Blender*, Feb. 2021, [Online]. Available: https://github.com/federicoarenasl/Data-Generation-with-Blender

[4] G. Jocher *et al.*, *ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation*, version v7.0, Nov. 2022. DOI: 10.5281/zenodo.7347926.

[5] Riis, Aerial sheep dataset, https://universe.roboflow.com/riis/aerial-sheep, Open Source Dataset, visited on 2023-12-04, 2022. [Online]. Available: https://universe.roboflow.com/riis/aerial-sheep.

[6] B. Dwyer *et al.* Roboflow (Version 1.0) [Software]. Available from https://roboflow.com. computer vision.