AUTONOMOUS UAV-BASED CATTLE DETECTION AND COUNTING USING YOLOV3 AND DEEP SORT

Muhammad Muaz Bin Abni Hajar [i] , Izzuddin Mat Lazim [ii], Ahmad Rasdan bin Rosdi [i] Liyana Ramli [i]

[i]. Department of Electrical and Electronic Engineering, Faculty of Engineering and Built Environment, Universiti Sains Islam Malaysia.

[ii] (*Corresponding author*) Department of Electrical and Electronic Engineering, Faculty of Engineering and Built Environment, Universiti Sains Islam Malaysia. nh.izzuddin@usim.edu.my

Abstract

Unmanned Aerial Vehicles (UAV) have been one of the most crucial technological advances in this era. One of the sectors that have been taken over gradually by UAV is agriculture. UAV has assisted the farmer in managing farmland autonomously especially in monitoring livestock. This study proposes the cattle detection and counting system based on deep learning via UAV. This detection and counting system is done autonomously for monitoring livestock, specifically cattle. We take advantage also of the assistant of UAVs to replace the "human eye" or real-time image in monitoring the livestock. The deep learning method used in this research is the YOLOv3 and Deep SORT that have been integrated into drone real-time images to capture the real-time cattle monitoring system.

Keywords: *UAV, YOLOv3, Deep SORT, Real-time*

INTRODUCTION

The agriculture business has evolved to meet the demands of humans, particularly in cattle production. Due to the depopulation of rural areas, agriculture has numerous economic issues as well as a labour shortage. Also, there is a growing trend toward animal consumption but less in production due to migration. Therefore, minimizing the problem by modern livestock approaches needs to consider human health while progressively improve the farm quality in the agriculture industry (Daponte et al., 2019). Modern monitoring techniques can improve animal output and quality, as well as their health. Data from livestock monitoring can be taken for analysis, and the farmer can decide without being involved in the process directly. Monitoring in modern agriculture is effective in determining levels of activity, health, food intake, and water consumption (Ismail et al., 2019). Conventional monitoring is manually done which the help of farmers to identify livestock by labelling those animals using ear tags, ear notch, tattoo, or marks.

The focus of this study is on object detection and counting methods for livestock monitoring using deep learning. The method is appropriate for such application because we all know that farming places are covered up on huge property, making physical inspection difficult. Deep learning, a subtype of artificial intelligence, will be discussed extensively in this study. The advancement of object detection is making more accurate detection and data for the image or video format in the so-called neural network. The artificial neural network is like a human brain that can link every neuron to thousand or millions of other neurons for interpreting the images. Better CNN algorithms and architecture is developed since 2016 such as real-time object detection.

YOLO uses a neural network to predict bounding boxes and their probabilities. This makes the YOLO one of the fastest and efficient ways to implement the livestock monitoring system in the world (Benjdira et al., 2018). Hence, A new study has proposed how to integrate an autonomous drone with a livestock monitoring system. The idea is to use the DJI Tello drone, which can fly autonomously at a great height and relay the information back to the farmers on the ground without intervention during the flight. Integration of this autonomously operated drone with the proposed livestock recognition and counting is something novel and may benefit the farmers. Object detection and counting can be implemented in the automation of livestock monitoring. Hence, this study will introduce how to integrate between object detection and counting system while implementing on a drone Tello for livestock monitoring.

METHODOLOGY

The method is based on a recent research study that can be enhanced and put into practice. The YOLOv3 real-time object detection system (Deep SORT) is used in this study to improve the data of a livestock monitoring system using deep learning. The proposed method focuses on the software part which is using python language. However, this study proposes the most recent approach of object detection, which employs the YOLOv3 real-time object detection system's neural network. This method differs from object detection methods such as thresholding and a deep learning variant known as convolution neural network (CNN) (Sharma & Mir, 2020).

The project aims to identify and count the types of livestock especially the cattle which are operated using drones. The drone can be controlled remotely from a distance or autonomously generate the pathway. It is one of the most affordable and portable drones available named DJI Tello Drone. Thus, this research proposing the method of the YOLOv3 version and Deep SORT with the help of drone application for autonomous path and counting.

Object Detection

Object detectors are novel machine learning models that were only made viable by breakthroughs in the field that gave rise to deep learning. Because the YOLO project is real-time, accuracy is equally as critical as processing speed and power efficiency. The first phase is object recognition. An object like cattle needs to be detected first before the system starts the counting. This method requires the YOLOv3 algorithm while linking to the python library named OpenCV. The most recent version of the YOLO detection model increases accuracy by sacrificing memory and time to some degree. Version three outperformed many of its peers, including several CNN models from the multi-stage camp, which had previously beaten its inaccuracy.

YOLO Architecture

The YOLOv3 algorithm is a multi-stage, single-shot detector that does localization and classification at the same time. The most recent version is more complex, with several more layers to load into the main memory than in previous versions of the algorithm and as a result, each pass through the networks is longer than in previous versions. The ultimate target of a single-shot detector that does localization and classification at the same time, the anchor box approach to bounding box regression, non-max suppression of IOUs, and augmented classification training were all included in this latest iteration from its predecessor (Scholarworks & Ordonia, 2019).

In terms of machine learning, YOLOv3 prediction on the object is using logistic regression which is supervised learning. YOLOv3 is extracting the data feature into a 53-convolution layer which is then named Darknet-53 as the backbone. The YOLOv3 is compared with the older version YOLOv2 (Redmon & Farhadi, 2018). The newer YOLOv3 model could allow detection at multiple levels of upsampling thanks to the deeper Darknet-53 feature extraction component. The technique of raising the size of an image is known as upsampling. Smaller artefacts can be identified more quickly because of the increased number of detection cells as in Figure 11.

In addition, darknet-54 has 54 convolution layers that resize an image from W x H to prediction maps of three different scales. The network is fed with the image of the dog in figure 3.1. Multiple convolutions are applied to the image, each using leaky ReLU and batch normalization. They use leaky RELUs, as well as several other leaky activation features. Leaky ReLU is a better version for ReLU as the activation function for many types of neural networks used today.

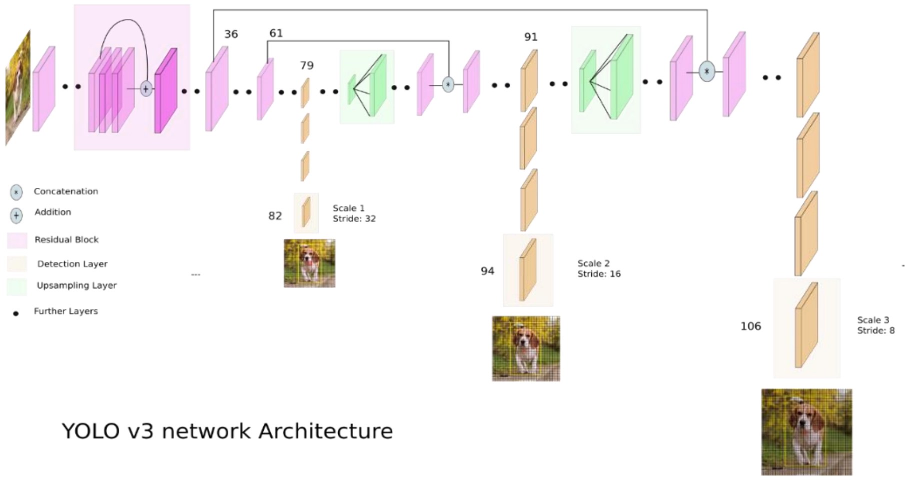

Figure 11. Yolov3 Architecture (Scholarworks & Ordonia, 2019)

Dataset

For the dataset, YOLOv3 is using the pretrained MS COCO from Microsoft. The COCO format for storing annotations is widely utilized when creating a new custom object detection dataset due to the dataset's popularity. The COCO dataset also includes annotations for tasks such as segmentation, allowing users to train new objects. MS COCO has a large amount of data for object detection by consisting of more than 300000 images and more than 80 categories of class prediction (Chen & Miao, 2020). All networks use YOLOv3 weights that are the component for pretrained data on the MS COCO dataset to be used in object detection.

Object Counting

The second phase in livestock monitoring is counting. This study proposed the method from YOLOv3 but applying other methods for specialized in tracking multiple objects which are Deep Simple Online and Real-time Tracking (Deep SORT). The proposed method is adequately better for continuing object detection from YOLOv3. Thus, this study focuses on how to count the number of cattle, especially in the farmland using deep learning of YOLOv3 and Deep Sort.

Multiple Object Tracking

Object tracking must be considered before counting the objects that have been identified. The object should be monitored so that after the frame is moved, the new object detection of the same object will be the same. As a means of locating many items in a sequence of frames, tracking is simple to comprehend (video). Multi-Target Tracking is a broad concept that encompasses a variety of tracking scenarios and most of the research found is tracking people (Bathija, 2019), automobiles (Santos et al., 2020), livestock, and other objects that may be used as targets. The MOT can be performed in two steps: Detection and Association (Santos et al., 2020). This project used the YOLOv3 method for detection. Next, the step for object tracking is association data which is used for every bounding box. Each target's bounding box geometry is estimated by predicting its new position in the most recent frame while assigning detections to the existing targets.

Tracking and counting using Deep SORT

Deep SORT improves tracking performance by adding apparent feature information matching. This enhancement allows the system to track the target throughout a longer duration of occlusion, reducing the number of ID transformations required (Jie et al., 2021). Furthermore, the thing to be considered also in the project is Non-Maximum Suppression (NMS). NMS is a specific targeting detection algorithm. During model prediction, it is employed for the secondary screening detection frame. If the intersection ratio of the candidate box and the box with the highest current target score is greater than or equal to the threshold, the threshold was set. The candidate boxes will be eliminated immediately if it is greater than or equal to the threshold. This method is far too straightforward, and it is prone to false and missing detection. When two objects are covered by each other, the NMS method may easily screen out the prediction box of the covered object (Jie et al., 2021).

Unmanned Aerial Vehicle (UAV)

The Unmanned aerial vehicle (UAV) has many applications such as surveillance, search and rescue, as well as crop seeding and spraying (Mat Lazim et al., 2021). In this study, the final step in completing the livestock monitoring project is to use the drone to implement the neural network from the previous method. It will be difficult for the farmer to count the livestock because he would have to wander around the farmland, especially if it is a huge region. The drone is capable of sensing objects automatically.

Furthermore, DJI Tello as shown in

Figure 12 is popular for educational use due to the availability of a software development kit (SDK). The video stream can record up to HD 720p in 30Fps. The method of using the drone is like an extension from a webcam video and provides the ability to capture photos or video from the drone's camera's perspective. The camera can be used to simulate flights or programs controlled by a machine learning model.

Figure 12. DJI Tello Drone EDU version

RESULT

Object Tracking and Counting

Object detection is run using PyCharm IDE which consists of the YOLOv3 algorithm and weights for pretrained dataset MSCOCO. The datasets have 30 class predictions including cows and people. The YOLO algorithm is tested to establish a fair way to evaluate the performance of the dataset. We have carried out the object detection model on MSCOCO datasets with two different labels. Each of their composition and results obtained are discussed below. Besides, the initial goal was to benchmark how well the YOLO models could track and count cattle using Deep Sort with the integration of the drone in real-time. The implementation is then will be discussed on the below subtopic with the result obtained through this project.

Object Tracking using YOLOv3

As a preliminary study, this research has used the dataset for people to replace the cow dataset. Hence, the initiative is taken by replacing the object detection with people as shown in Figure 13. The detection is successfully done with the correct dataset and label of the people.

Figure 13. Human and cow detection

Livestock Counting Using Deep SORT

The result for the implementation of object detection using YOLOv3 and the Deep SORT tracking which by labelling of every bounding box detected on people is shown in Figure 14. This is one example of simple object tracking by counting the unique or multiple bounding boxes appear in a different time and places. The object detected with YOLOv3 and Deep Sort enables counting to detect more boxes and achieve a higher average confidence value which makes the counting process is effective for a real-time detection system.

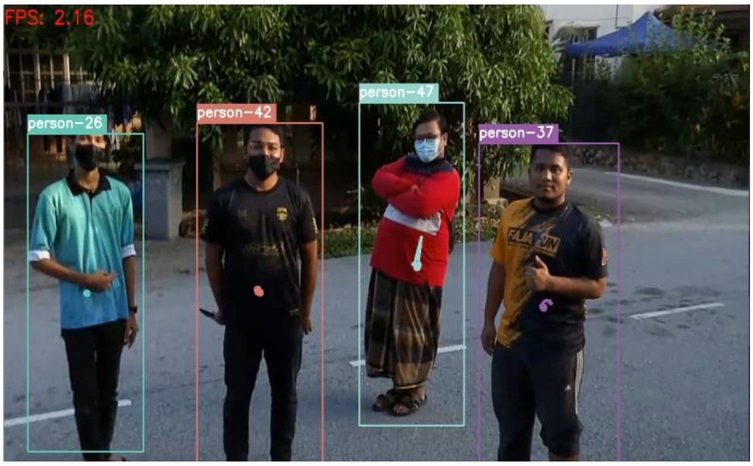

Figure 14. Human Tracking using Tiny-YOLOv3

Furthermore, the concern for this project also is about the fps used during real-time tracking algorithm running. As to be mentioned, there is two different object detection system used in this project. At first, the project is focused to use the object detection of YOLO version 3. Unfortunately, after some of the program execution with the Deep SORT algorithm, the fps is dropping abruptly into 0.4 to 0.5 fps. This happens because the YOLOv3 with the addition of Deep SORT is required a lot of higher performance. In addition, the performance of these methods comes at increased computational and implementation complexity compared to SORT only which is a much simpler framework. This simple approach achieves favorable performance at high frame rates (Wojke et al., n.d.).

Hence, the detection system is changed into a simpler version called Tiny-YOLOv3 which is faster. The Tiny-YOLO architecture is approximately 442% faster than the larger version of YOLOv3, achieving upwards of 244 FPS on a single GPU (Adarsh et al., 2020). The small model size which is less than 100MB and fast inference speed make the Tiny-YOLOv3 object detector naturally suited for embedded computer vision devices like the DJI Tello drone.

Autonomous Flight Path

DJI Tello UAV has a built-in programmable control command code. The command code is set in the text file so that the drone can read and execute the line of command before running the Deep SORT algorithm. The drone's flight path is autonomously set using the python library called Tkinter. Figure 15 shows the GUI from Tkinter which consists of many features. As the button is clicked, the data is written in the textbox to indicate the command for the drone. After finish and exit the Tello GUI, the data will be updated into the text file. The text file will be used during the Deep SORT program to assign the command to the Tello drone.

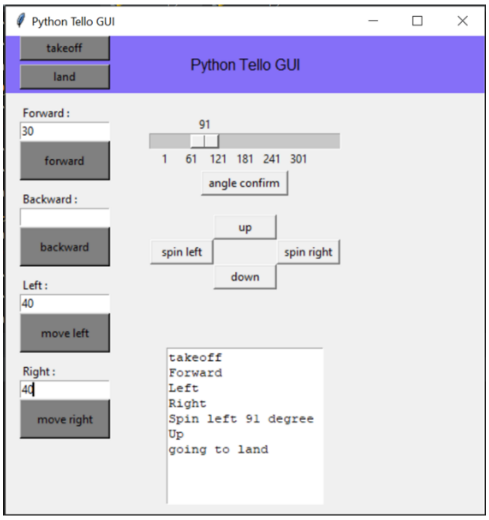

Figure 15. GUI Path Plan

The text file is read in the drone. The drone is programmable towards the command given and set in the user guide. Thus, by providing the path using Tello GUI, the drone can move accordingly based on the path given.

Integration of the Tracking Deep SORT into Drone

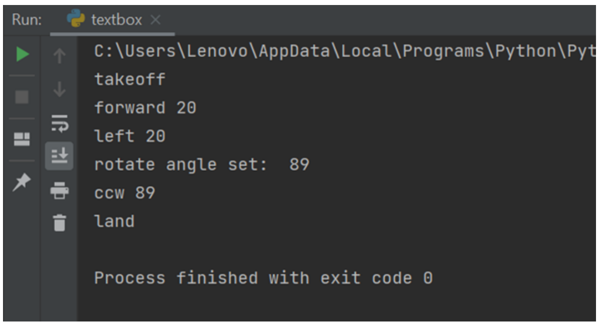

Figure 16. Terminal command sent to the UAV

The drone connection is based on the Djitellopy library and can be found in the PyCharm IDE library download. Figure 6 shows the result of the successful connection through terminal Pycharm. There will be two different programs for a single running code file which are the tracking, and the path planning. Each program is divided by the time slot given. Thus, by doing the function of the asynchronous library the program especially for drone autonomous path to execute each command is easier. The frame shown will be smoother when running those two programs together.

Figure 17 the result for the integration of the tracking cattle and drone path with the FPS at about 3 to 5 FPS.

Figure 17. Cattle tracking and Counting using Deep SORT

CONCLUSION

This paper presents the system for detecting and counting cattle using autonomous UAVs for livestock monitoring applications. The detection and count system are developed by using a deep learning approach through the Tiny-YOLOv3 model and Deep Sort. Meanwhile, the autonomous flight of the UAV is achieved by using a developed GUI. The results from a simulated environment show that the proposed system is successfully implemented. For future work, a more capable UAV can be considered for better tracking and detection of the cattle.

ACKNOWLEDGMENT

This research is supported by Universiti Sains Islam Malaysia (USIM) through USIM Research Grant No. PPPI/USIM-RACER_0120/FKAB/051000/12920.

REFERENCES

Adarsh, P., Rathi, P., & Kumar, M. (2020). YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. *2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020*, 687–694. https://doi.org/10.1109/ICACCS48705.2020.9074315

Bathija, A. (2019). Visual Object Detection and Tracking using YOLO and SORT. *International Journal of Engineering Research and Technology (IJERT)*, *8*(11), 705–708. https://www.ijert.org

Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., & Ouni, K. (2018). Car detection using unmanned aerial vehicles: Comparison between faster R-CNN and YOLOv3. *Proceedings of the 1st International Conference on Unmanned Vehicle Systems (UVS)*, 1–6.

Chen, B., & Miao, X. (2020). Distribution Line Pole Detection and Counting Based on YOLO Using UAV Inspection Line Video. *Journal of Electrical Engineering and Technology*, *15*(1), 441–448. https://doi.org/10.1007/s42835-019-00230-w

Daponte, P., De Vito, L., Glielmo, L., Iannelli, L., Liuzza, D., Picariello, F., & Silano, G. (2019). A review on the use of drones for precision agriculture. *IOP Conference Series: Earth and Environmental Science*, *275*(1). https://doi.org/10.1088/1755-1315/275/1/012022

Ismail, Z. H., Chun, A. K. ng K., & Shapiai Razak, M. I. (2019). Efficient Herd - Outlier Detection in Livestock Monitoring System Based on Density - Based Spatial Clustering. *IEEE Access*, *7*, 175062–175070. https://doi.org/10.1109/ACCESS.2019.2952912

Jie, Y., Leonidas, L., Mumtaz, F., & Ali, M. (2021). Ship detection and tracking in inland waterways using improved yolov3 and deep sort. *Symmetry*, *13*(2), 1–19. https://doi.org/10.3390/sym13020308

Mat Lazim, I., Husain, A. R., Mohamed, Z., Mohd Basri, M. A., Mohd Subha, N. A., & Ramli, L. (2021). Effective Formation Tracking of Quadrotors with Intelligent Disturbance Observer-Based Control. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering 2021 45:3*, *45*(3), 761–776. https://doi.org/10.1007/S40998-021-00417-W

Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *Tech Report*, 1–6. https://pjreddie.com/media/files/papers/YOLOv3.pdf

Santos, A. M., Bastos-filho, C. J. A., Maciel, A. M. A., & Lima, E. (2020). Counting Vehicle with High-Precision in Brazilian Roads Using YOLOv3 and Deep SORT. *Polytechnic School of Pernambuco - University of Pernambuco Recife,*.

Scholarworks, S., & Ordonia, S. (2019). *Detecting Cars in a Parking Lot using Deep Learning*.

Sharma, V., & Mir, R. N. (2020). A comprehensive and systematic look up into deep learning based object detection techniques: A review. *Computer Science Review*, *38*, 100301. https://doi.org/10.1016/j.cosrev.2020.100301

Wojke, N., Bewley, A., & Paulus, D. (n.d.). *SIMPLE ONLINE AND REALTIME TRACKING WITH A DEEP ASSOCIATION METRIC Nicolai Wojke † , Alex Bewley , Dietrich Paulus † University of Koblenz-Landau † , Queensland University of Technology*.