Yuzuncu Yil University Journal of Agricultural Sciences, Volume: 33, Issue: 3, 30.09.2023

Yuzuncu Yil University
Journal of Agricultural Sciences
(Yüzüncü Yıl Üniversitesi Tarım Bilimleri Dergisi)
https://dergipark.org.tr/en/pub/yyutbd

ISSN: 1308-7576 e-ISSN: 1308-7584

Research Article

Effects of Data Augmentation Methods on YOLO v5s: Application of Deep Learning with Pytorch for Individual Cattle Identification

Cafer Tayyar BATİ*[1], Gazel SER[2]

Van Yuzuncu Yil University, Faculty of Agriculture, Department of Animal Science, Van, Turkey

[1]https://orcid.org/0000-0002-4218-4974, [2]https://orcid.org/0000-0003-2756-4116

*Corresponding author: cafertayyarbati@gmail.com

Article Info

Received: 02.02.2023
Accepted: 23.05.2023
Online published: 15.09.2023
DOI: 10.29133/yyutbd.1246901

Keywords

Cattle,
Data augmentation,
Object detection,
YOLO

Abstract: In this paper, we investigate the performance of the YOLO v5s (You Only Look Once) model for the identification of individual cattle in a cattle herd. The model is a popular method for real-time object detection, accuracy, and speed. However, since the videos obtained from the cattle herd consist of free space images, the number of frames in the data is unbalanced. This negatively affects the performance of the YOLOv5 model. First, we investigate the model performance on the unbalanced initial dataset obtained from raw images, then we stabilize the initial dataset using some data augmentation methods and obtain the model performance. Finally, we built the target detection model and achieved excellent model performance with an mAP (mean average precision) of 99.5% on the balanced dataset compared to the model on the unbalanced data (mAP of 95.8%). The experimental results show that YOLO v5s has a good potential for automatic cattle identification, but with the use of data augmentation methods, superior performance can be obtained from the model.

To Cite: Bati, C T, Ser, G, 2023. Effects of Data Augmentation Methods on YOLO v5s: Application of Deep Learning with Pytorch for Individual Cattle Identification. *Yuzuncu Yil University Journal of Agricultural Sciences*, 33(3): 363-376.
DOI: https://doi.org/10.29133/yyutbd.1246901

1. Introduction

The increasing demand for animal products requires breeders to increase production without compromising animal welfare (Masebo et al., 2023). Therefore, it is important to solve on/off-farm management problems with less cost. Livestock farmers and producers are very concerned with the identification of their animals for management simplicity. It is known that most farms today use artificial farming methods and many of them utilize ear tags to count or identify the animals (Zhang et al., 2022). Traditional techniques for animal identification other than ear tags include microchips and Radio Frequency Identification (RFID) tags. All of these methods require direct contact and can cause varying degrees of injury to animals. Animal recognition and well-being evaluation utilizing non-invasive digital technology has received attention in agriculture in recent years, particularly for accuracy tracking (Dac et al., 2022; Zhang et al., 2022). Recently, a lot of work has been done in livestock management, including modern and intelligent methods using computer vision methods utilizing deep learning. Image processing and object detection have become very popular in these studies with the advancing graphics processing unit (GPU) technology. Object detection studies with computer vision are quite remarkable and have been extensively used in animal identification. Features such as body patterns, muzzle structure and posture are used to recognize individual farm animals. R-CNN object detection has been used to detect the tail of a cow for body condition score (Huang et al., 2019) and to identify individual cattle (Andrew et al., 2017). Among the object detection methods, YOLO (You Only Look Once) is a remarkable and useful tool. The YOLO v5 algorithm has been used by different researchers for accurate identification of domestic cattle (Luo et al., 2022), cattle counting from UAV images (de Lima Weber et al., 2023), face identification in dairy cows (Dac et al., 2022), accurate and fast detection of goats (Zhang et al., 2022), and sheep behavior identification (Chen et al., 2022). Recent studies have shown that deep learning technologies that are non-invasive and do not affect animal welfare have the potential to provide ease of livestock identification and management (Bati and Ser, 2023; Subedi et al., 2023). In deep learning studies, unbalanced datasets are encountered, which affect network performance and are a chronic problem for almost every study. Especially in image classification or object detection studies in animal production, it is challenging to get the same amount of data such as images and videos for each class.

This study was designed to investigate the performance loss caused by unbalanced datasets in computer vision studies and to provide solutions. The rest of the paper, in which we use object detection from Holstein Friesian cattle images for individual cattle identification and improve the performance of the method with data augmentation methods, is arranged as follows; In the section two, the original dataset, the methodology used and experimental setup are presented. The third section presents the individual cattle identification results and the last two sections present the discussion and conclusion respectively.

2. Material and Methods

2.1. Original dataset

In this study, we used the FriesianCattle2017 (Andrew et al., 2016; Andrew et al., 2017) dataset, which consists of 940 RGB images of 89 different Holstein Friesian cattle. The data was captured in a real closed farm context, over a two-hour session, using a camera mounted on a walkway between pens and milking stations (Andrew et al., 2017). More detailed information on dataset acquisition can be found in Andrew et al. (2016) and Andrew et al. (2017). Sample frames of individuals in the dataset are presented in Figure 1.

Figure 1. Sample frames from the original dataset.

2.2. Deep neural architectures: YOLO v5s object detection algorithm

The YOLOv5 network model is one of the state-of-the-art object detection algorithms with high detection accuracy proposed by Glenn Jocher in 2020 (Jocher, 2020). The YOLO v5 network has the highest computational speed because the size of the weight file of the target detection network model is small (Chen et al., 2022). This shows YOLO v5 is appropriate for applying real-time detection. We preferred the YOLO v5s network for cattle identification due to its advantages such as high detection accuracy, lightweight features, as well as detection speed (Yan et al., 2021). YOLO v5 network consists of four main components, namely the input, backbone (CSPDark net), neck(PANet), and head. Input terminal mainly includes data preprocessing, including mosaic data augmentation and adaptive image filling (Li et al., 2022). Once the image is input, it is aggregated in the backbone and generates image features at different image detail levels. It was based on the Cross-Stage-Partial-Network (CSPNet) (Jintasuttisak et al., 2022). The neck then combines the image features and passes them to the prediction layer. YOLO v5 uses the Path-Aggregation-Network (PANet) (Liu et al., 2018) in the model neck for extracting feature pyramids. The head estimates image features to create bounding boxes and predictive category (Chen et al., 2022; Jintasuttisak et al., 2022). The architecture of the YOLO v5 is presented in Figure 2.

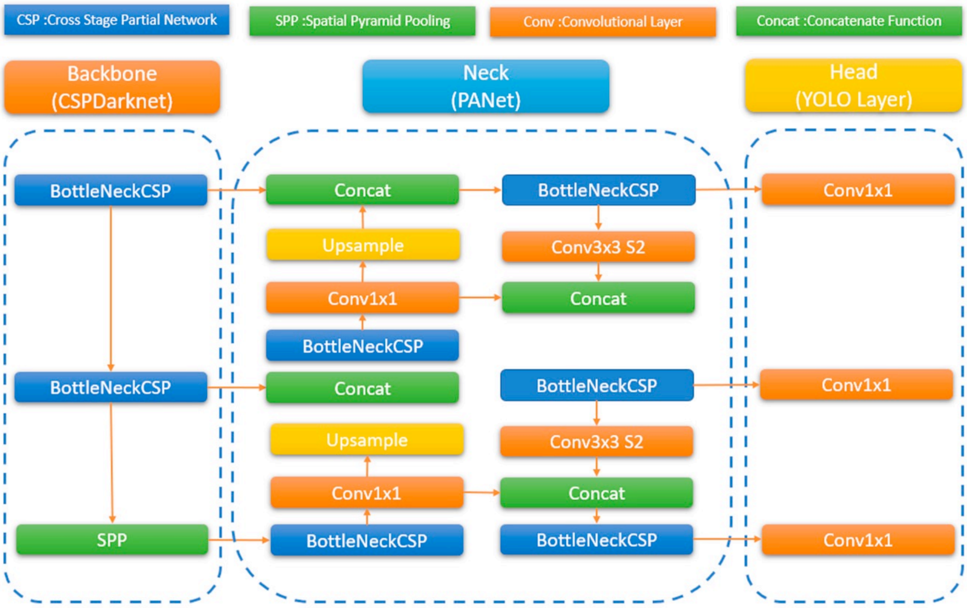

Figure 2. Architecture of the YOLO v5 (Egi et al., 2022).

2.2.1. Intersection over union (IoU)

This metric is used to find the match between each cattle's ground truth annotations and predicted bounding boxes to make a quantitative comparison of performance. As shown in Figure 3, we calculated the IoU as the area of the intersections between the predicted cattle bounding box and the ground truth divided by the area of the unions. The estimated bounding box with an IoU equal to or greater than 0.5 was considered the correct prediction of cattle. To evaluate the performance of the models comparatively, the number of accurate predictions was used in the calculation of performance metrics (Jintasuttisak et al., 2022).

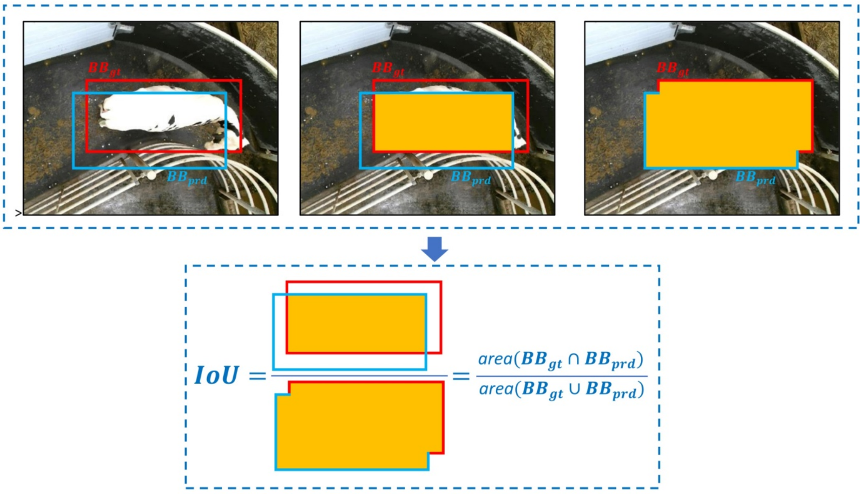

Figure 3. Calculation of IoU using predicted (BB_{prd}) and ground-truth bounding boxes (BB_{gt}). *Area* ($BB_{gt} \cap BB_{prd}$); Intersection of BB_{prd} and BB_{gt} *area* ($BB_{gt} \cup BB_{prd}$); Union of BB_{prd} and BB_{gt}.

2.3. Experimental setup

In this study, we followed the stages shown in Figure 4 for preprocessing, model training, and evaluation. In the first stage, we performed dataset creation (detailed description in section 2.3.1). In the second stage, we labeled and split the dataset (for training and validation). In the third stage, the mosaic data augmentation method was activated for YOLOv5, and training was performed. In the last stage, testing of the obtained models was performed.

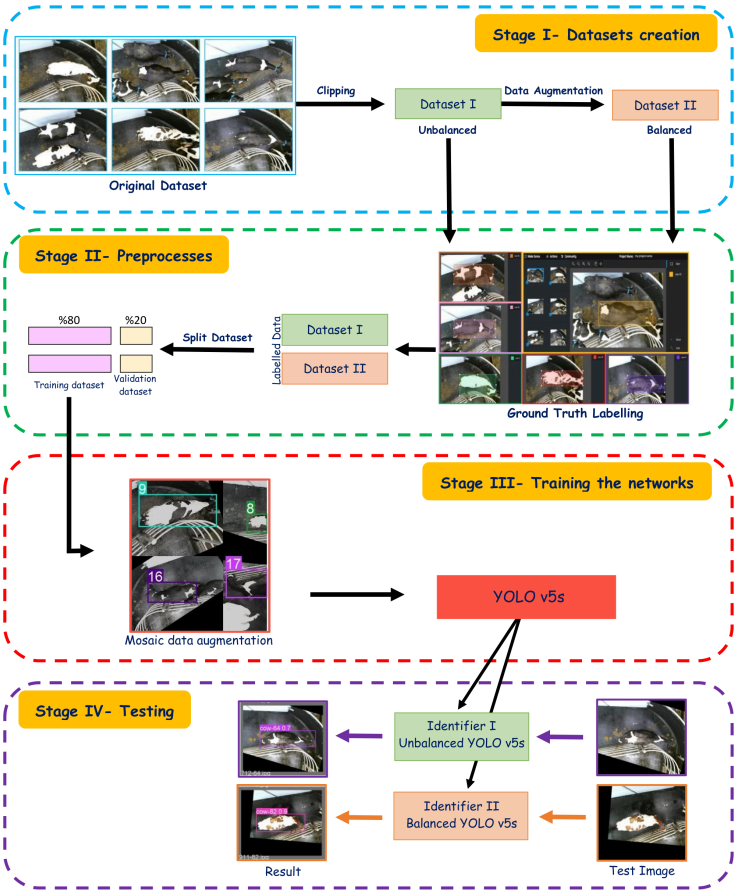

Figure 4. Data evaluation stages.

2.3.1. Experimental datasets

Considering the number of frames in Figure 5 for 89 different cattle in the original dataset, we used 20 Holstein Fresian cattle with 15 or more frames in our study and defined this dataset as Dataset I. Then, we created Dataset II by modifying Dataset I. A visualization showing the number of frames in both datasets is given in Figure 6.

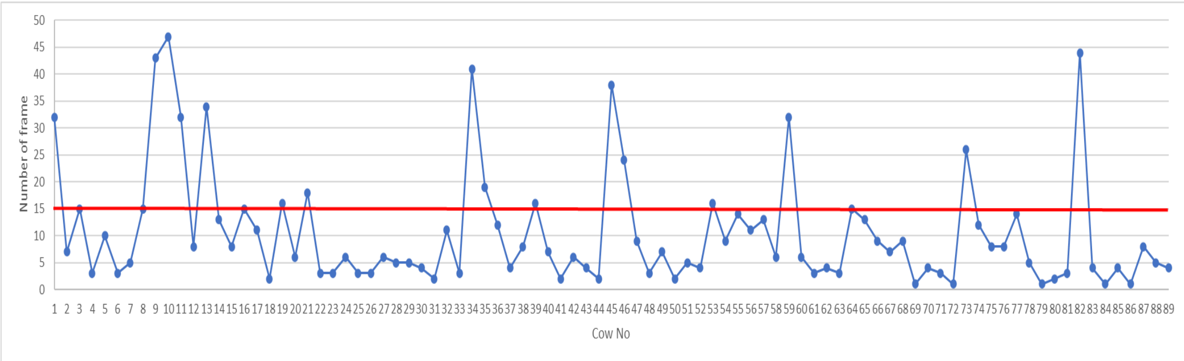

Figure 5. Number of frames in the original dataset.

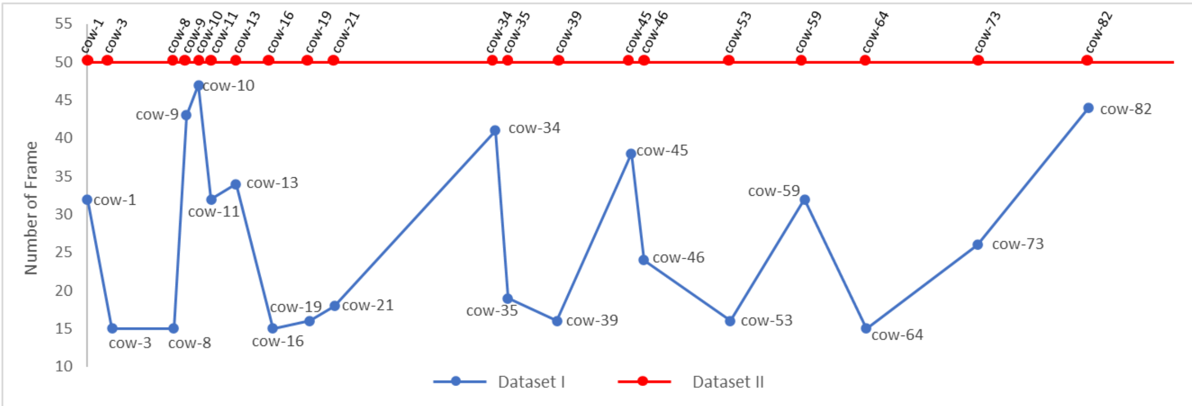

Figure 6. Number of frames in both datasets used in the study.

Detailed information on the construction of both datasets is provided.

Dataset I (Unbalanced): In this dataset, 539 images from videos of 20 cattle were used. While 432 of these images were used to train the models, 107 were used for testing. Figure 6 shows that the number of frames between classes is highly unbalanced due to the nature of the dataset. This is due to the fact that the animals were released on the walkway without any direction when the images were taken.

Dataset II (Balanced): In order to enrich the image data of the training set and to overcome the unbalance in Dataset I, we performed data augmentation using the "Augmentor" option in python. We used the data augmentation methods of rotate, zoom, skew, distortion, shear, and flip to arrange 50 frames in all classes of Dataset I. Thus, we balanced the dataset and used 800 frames for training and 200 frames for testing the models. Figure 7 shows examples of the application of data augmentation methods for Dataset I.

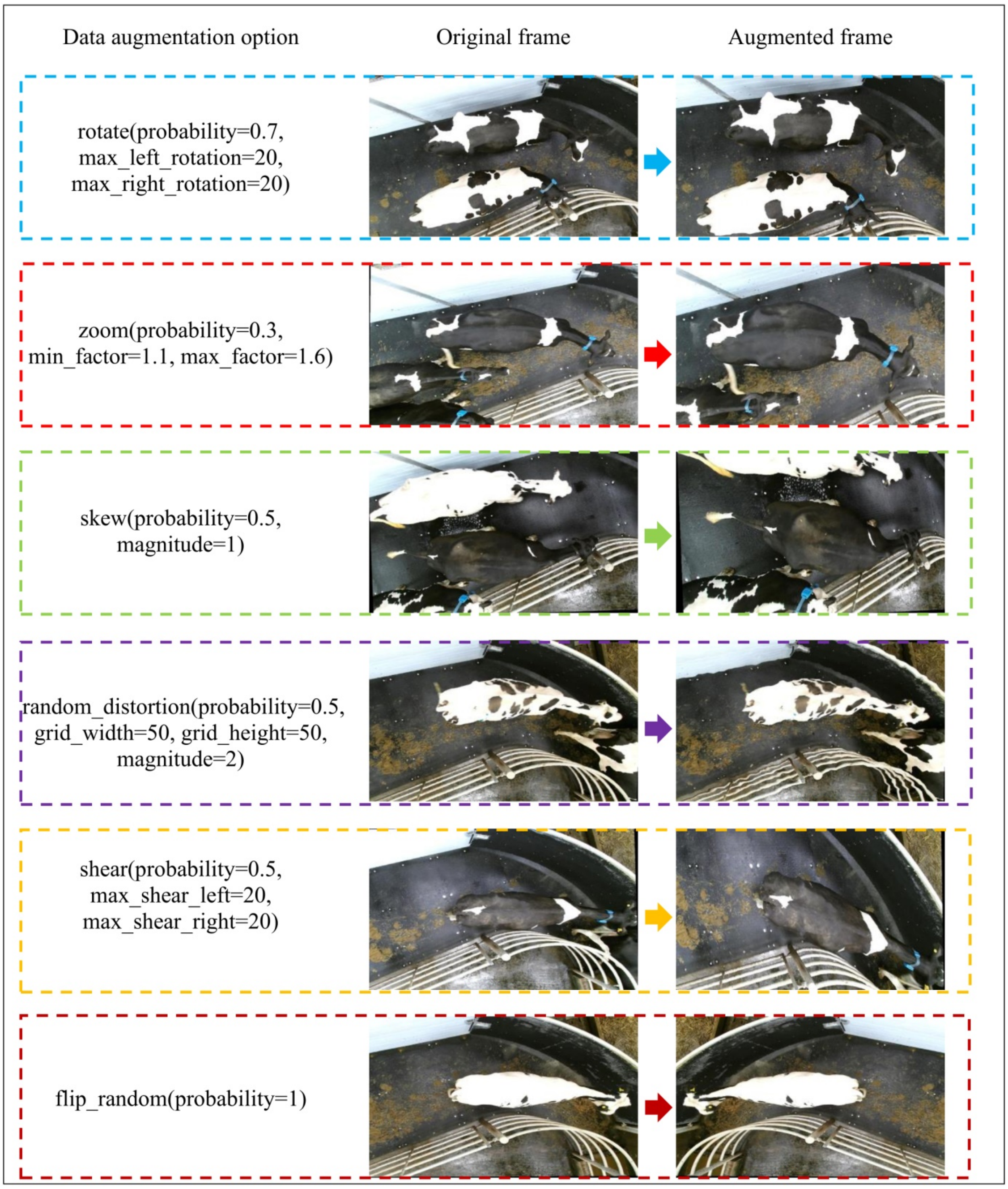

Figure 7. Implementation of data augmentation methods on the dataset I.

2.3.2. Ground truth labeling

In order to manually annotate cattle in the images, rectangular boxes were drawn to include the target cattle using the MakeSense (Skalski, 2019) image data annotation tool. The labeling process was completed by assigning cattle numbers to the drawn boxes. Afterward, the label package was created by saving the YOLO format files in zip format. The labeling process is given in Figure 8.

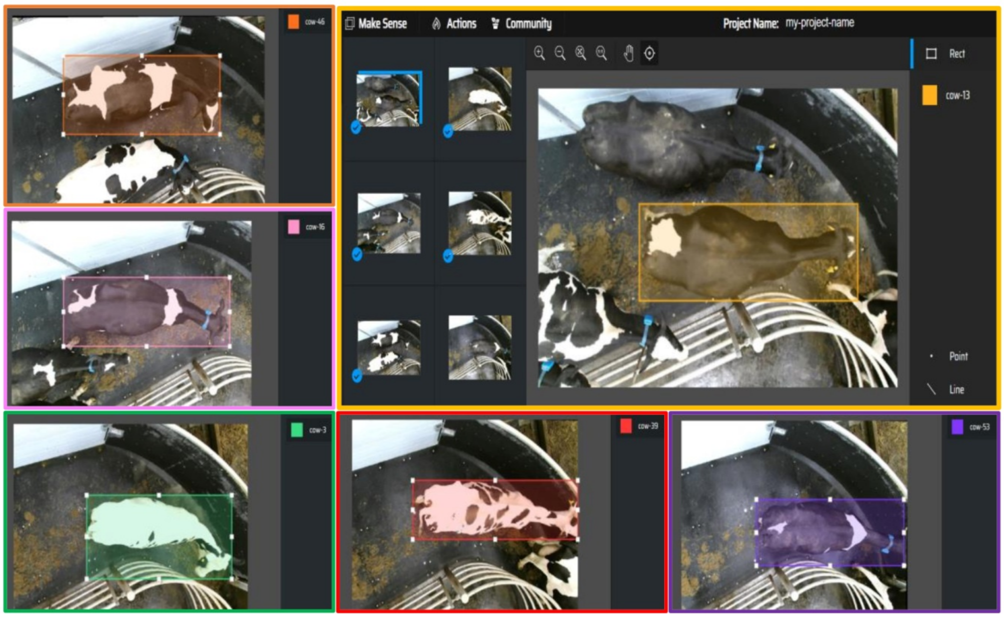

Figure 8. Ground truth labelling.

2.3.3. Network training parameters

In this paper, a ratio of 80%-20% is used for training and testing. The models have a batch size of 16, learning rate of 0.01, momentum of 0.937, epoch of 100, image size of 416, and optimization method of SGD. The open-source Python 3.8 (Van Rossum and Drake, 2009) package program and Pytorch 1.11.0 (Paszke et al., 2019), a high-performance deep learning library for YOLO v5s, were used in deep learning analyses. The information about some of the specific configurations we used is given in Table 1.

Table 1. Environment settings created for experimental processes

	Parameters	**Configurations**
Software	Operation system	Windows 10
	Framework for deep learning	Pytorch 1.11.0
	Langue of programming	Python 3.8
	IDE	Spyder 4.1.4
	GPU accelerated environment	CUDA 11.0
Hardware	GPU	GeForce RTX 2070, 8 GB GDDR6 Dedicated VRAM
	CPU	Intel Core i7-9750H @ 32 GB DDR4 2666 MHz

2.3.4. Performance evaluation

The criteria used to evaluate the performances of the models are Precision, Recall, and mAP (mean average precision). These metrics are calculated with tp (true positives), fp (false positives) and fn (false negatives) values obtained from the confusion matrix by Equation 1-3.

$$Precision = \frac{tp}{tp + fp} \quad (1)$$

$$Recall = \frac{tp}{tp + fn} \quad (2)$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (3)$$

Where;
AP: Average precision

3. Results

This section presents the performance results of the YOLO v5s algorithm for dataset I (unbalanced) and dataset II (balanced).

3.1. Dataset I results

Figure 9 shows the performance graphs of the YOLO v5s obtained from the training and validation dataset for the unbalanced dataset. The loss and mAP curves converge towards the 100th epoch. At the same time, the mAP value reached its highest value of 0.958 at epoch 76.

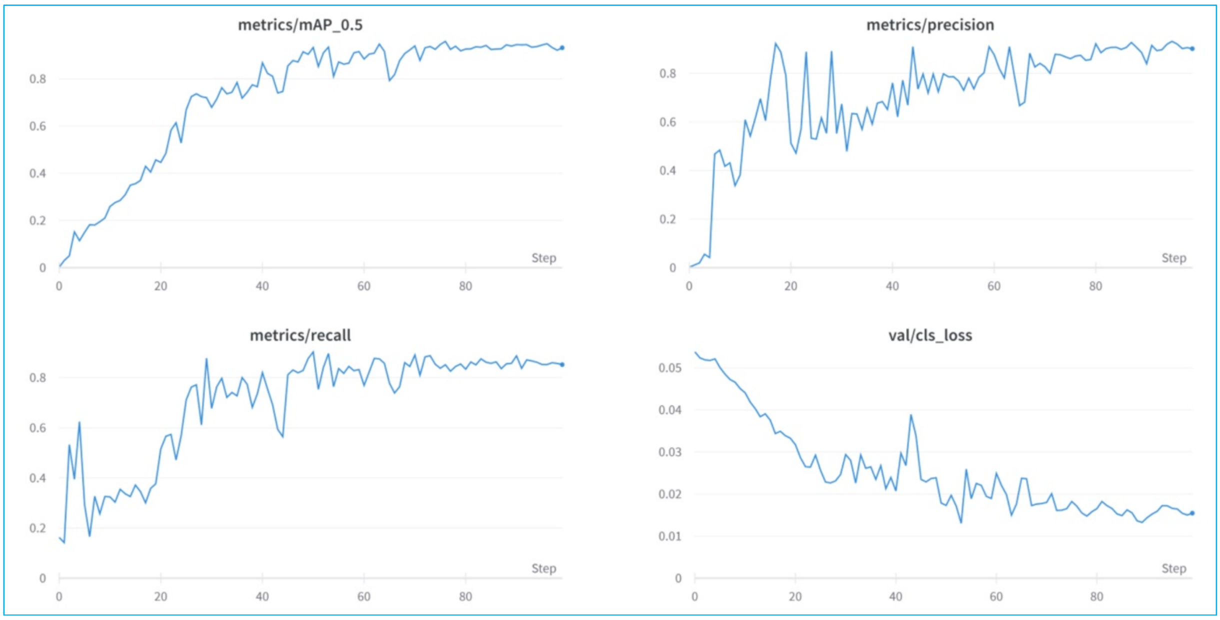

Figure 9. Model performance results trained on dataset I (Unbalanced).

The confusion matrix in Figure 10 demonstrates the functional ability of the YOLO v5s model to accurately predict the frames of cattle in images. The vertical and horizontal axes are the actual and predicted labels, respectively, and the diagonal represents the accurately identified frames. When the confusion matrix in Figure 10 is analyzed, all frames of 9 cattle are correctly identified. Only one of the 4 frames belonging to cow-35 was correctly identified, while one was identified as cow-39 and two were identified as background. It is understood that most of the frames that were actually cow-73 were identified as cow-53 and one frame that was actually cow-59 was identified as background. However, some frames that were actually cow-10, cow-19, cow-34 and background were predicted as cow-21. Images comparing some of the frames predicted by YOLO v5s with the actual labels are given in Figure 11.

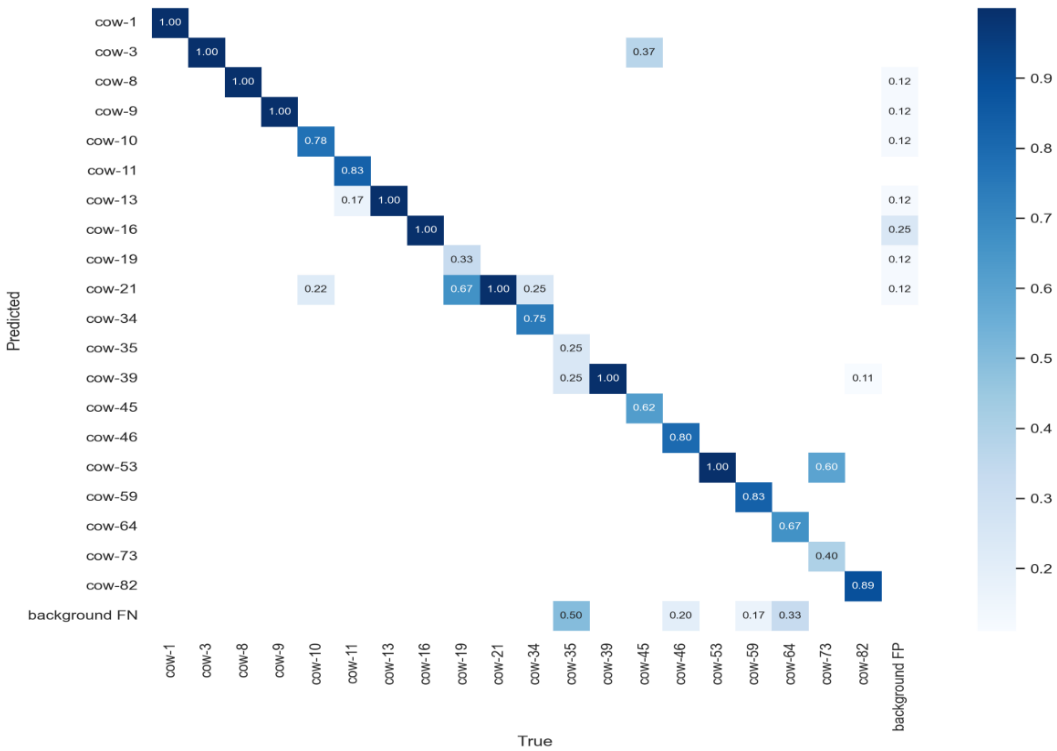

Figure 10. Confusion matrix of the model trained on dataset I (Unbalanced).

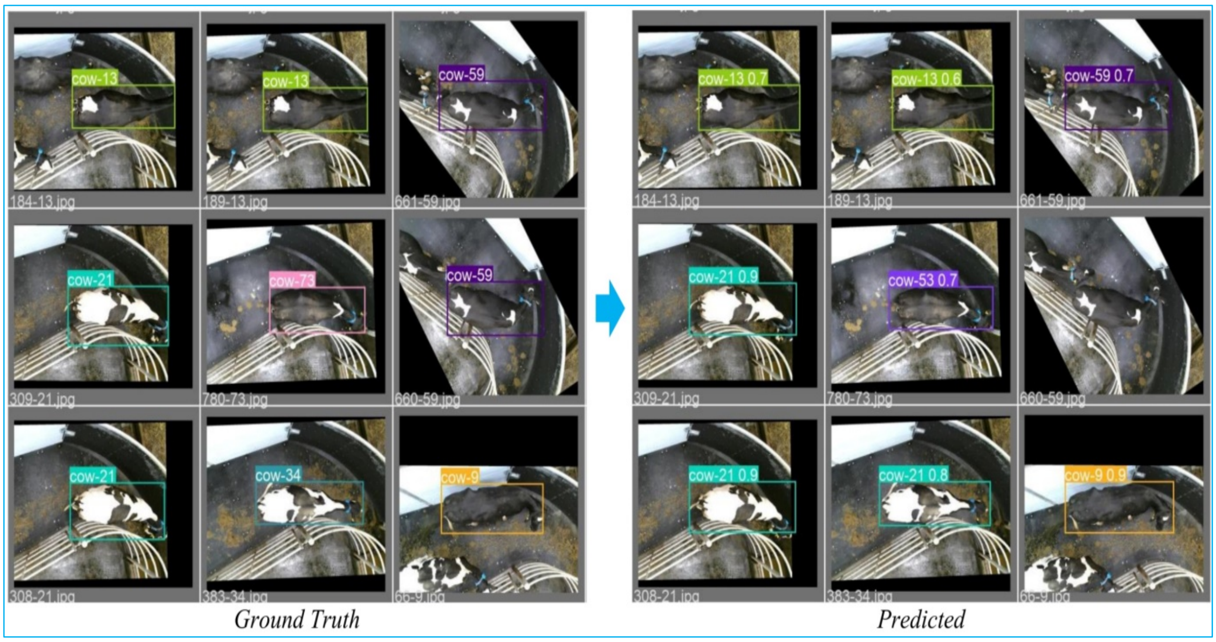

Figure 11. Comparison of ground truth and predicted frames for the model trained on the first dataset.

3.2. Dataset II results

The YOLO v5s model performance curves in the balanced dataset II are given in Figure 12. The model performance metric curves progressed smoothly. In contrast to the mAP curve in dataset I, here it reached high values after epoch 40 and reached its highest value of 0.995 at epoch 93.

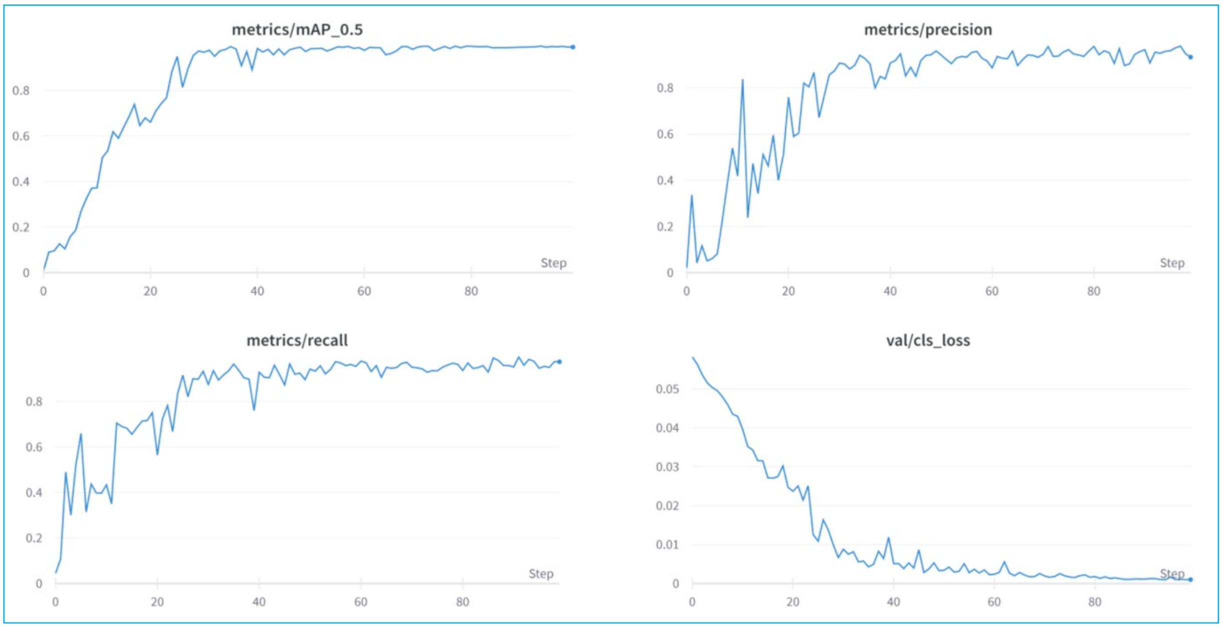

Figure 12. Model performance results trained on the second dataset (Balanced).

When the confusion matrix in Figure 13 is examined, it is understood that all images of all cattle except for only 3 cattle (cow-1, cow-46 and cow-59) are correctly identified. One image each of cow-1, cow-46 and cow-59 was identified as background. In addition, some of the images that were actually background (which could be an unlabeled cattle) were identified as cow-11 and cow-13. A comparison of some of the frames predicted by the model with the ground truth is shown in Figure 14.

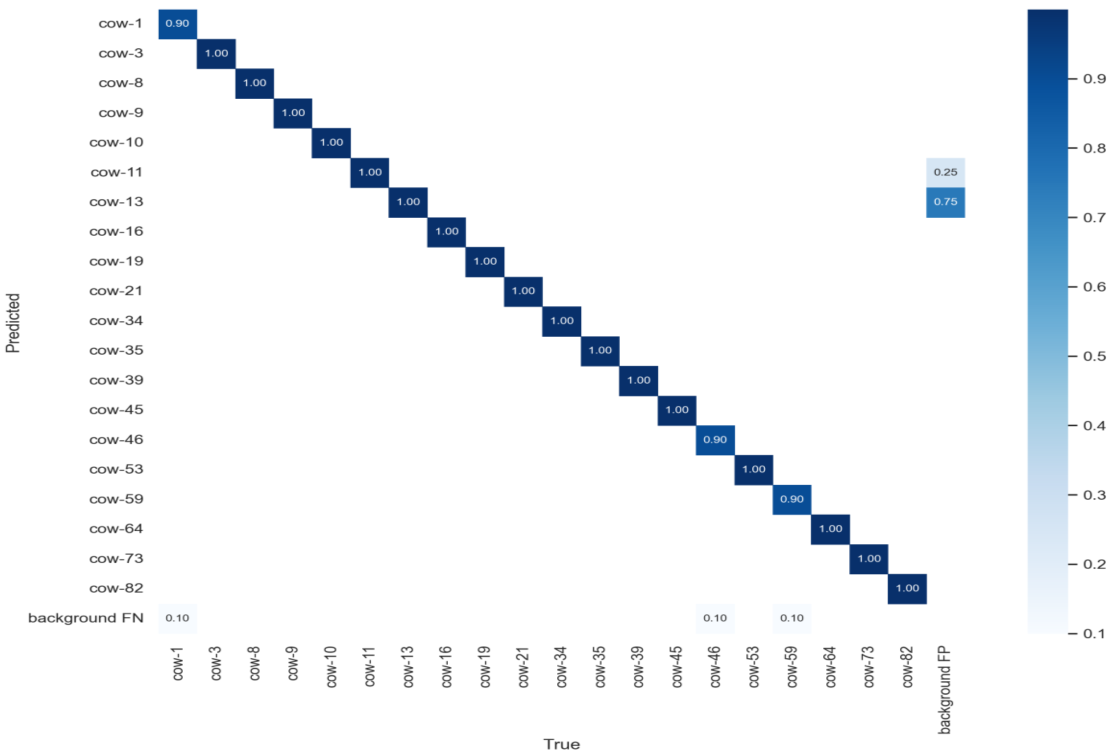

Figure 13. Confusion matrix of the model trained on dataset II (Balanced).

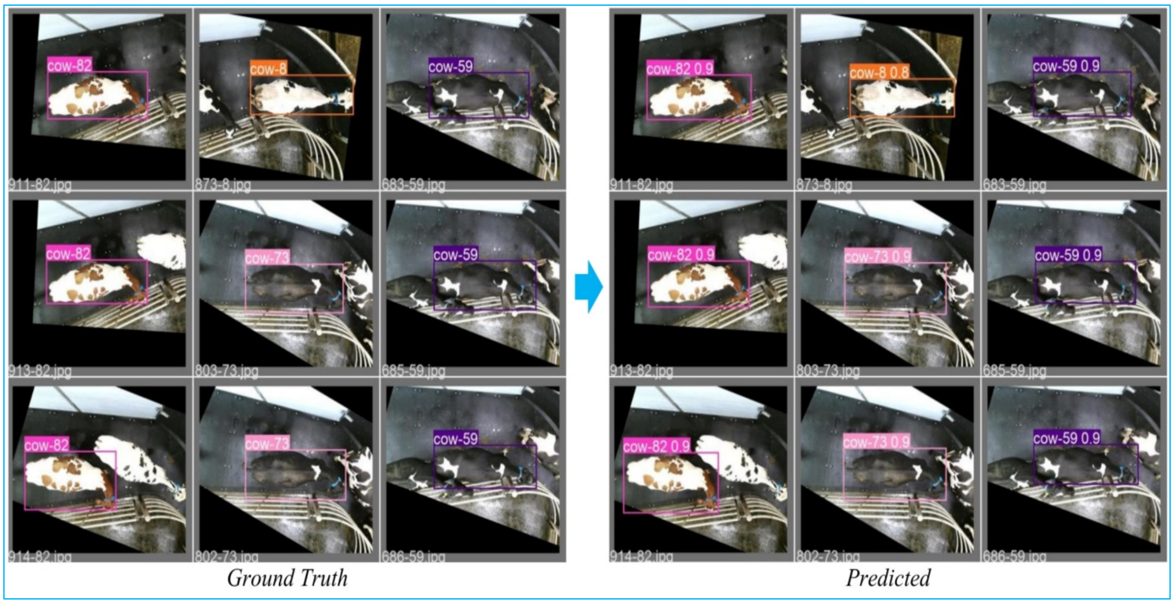

Figure 14. Comparison of ground truth and predicted frames for the model trained on the second dataset.

The performance metrics obtained from the YOLO V5s model to compare the classification accuracy for both dataset I and dataset II are given in Table 2. Balancing the dataset resulted in a significant increase in mAP (99.5%). In terms of recall and precision, balanced YOLO v5s gave the best performance. Comparing the training times, increasing the number of frames in the second dataset increased the training time by about 10 minutes.

Table 2. Individual cattle identification results

Networks	Training Time (mins)	Loss	Precision	Recall	mAP@0.5
Unbalanced YOLO v5s	16.467	0.015	0.871	0.851	0.958
Balanced YOLO v5s	26.600	0.001	0.950	0.985	0.995

4. Discussion

This study aims to determine the effect of data augmentation methods on YOLO v5s on an application dataset. For this purpose, the YOLO v5s object detection algorithm was applied to a dataset with unbalance in terms of the number of images between classes, and a second dataset was obtained by applying data augmentation methods and the results were discussed. Andrew et al. (2017), in their study on the original dataset of the study, achieved an mAP value of 0.993 in the cattle detection task and a mAP value of 0.861 in the individual cattle identification task with R-CNN networks. In the present study, we reduced the number of classes to 20 and obtained an mAP value of 0.995 with YOLO v5s in the dataset where we eliminated the unbalance.

In the present study, since the cattle identification performance obtained from the balanced dataset was approximately 100% (0.995 mAP), model training was performed with the default hyperparameters and no hyperparameter adjustments were made. However, previous studies on deep learning show that using various combinations of hyperparameters can result in architectures that are more suitable for the dataset (Ser and Bati, 2019; Altınbilek and Kızıl, 2022). Bocaj et al. (2020) used deep convolutional neural networks to identify the movements of horses and goats. In their study, they actively used hyperparameter tuning to optimize the overall accuracy. In another study, Zhang et al. (2022) developed a method based on YOLOv5 for goat head detection and automatic counting using combinations of data augmentation methods. In this study, they used combinations of mosaic, mixup, and RandAugment methods and it was reported that the most successful combination was obtained from YOLOv5 + Mixup + Mosaic + RandAugment. In addition, a convolutional neural network-based animal identification method for cattle and sheep was proposed by Sun et al. (2021). As in the present study, the original image data were firstly augmented by random cropping, random angle inversion, and random horizontal undoing. Then, a binary classification model for cattle and sheep recognition based on VGG-16 convolutional neural network is constructed. In addition, the relevant hyperparameters were constantly adjusted to increase the number of iterations in the study, and finally, a higher recognition accuracy rate was achieved.

Data augmentation provides an efficient way to extend training data and overcome the unbalance between classes (Shorten and Khoshgoftaar, 2019). However, the construction of models based on deep learning is realized on the basis of training a large amount of image data (Yan et al., 2021; Shojaeipour et al., 2021). The larger the size of the dataset and the better the data quality, the greater the ability of the model to generalize. Yet, while collecting data for real-life situations, classes in the data may not be equally represented, often due to conditions such as different environmental conditions. The unbalance between classes is known to affect network performance (Kasfi et al., 2016). Lee et al. (2022) used a YOLO network to monitor the invasion of wildlife animals on farms. In this study, they proposed a subtraction-addition data augmentation method, noting that creating a training dataset for specific wild animals requires considerable time and effort. In a comparison between the object detector trained using the proposed data augmentation technique and the object detector trained using existing data augmentation techniques, they found that the mAP increased by ≥2.2%. The mosaic data augmentation method, introduced by Bochkovskiy et al. (2020) in YOLO v4, uses four images from the training data and adds the training data to a single image through flipping, color gamut modification, and scaling (Zhang et al., 2022; Chen et al., 2022). In our study, we used mosaic data augmentation to expand the image set before feeding the images to the YOLO v5s network model. However, although the mosaic data augmentation method has advantages such as enriching the object detection back-ground and tiny objects and reducing the dependency on the batch size, it is unable to overcome the unbalance between classes. Therefore, in this study, the images in the first dataset were also included in the data augmentation process and then fed into the YOLO v5s network (Figure 5). According to the experimental results, the balanced classes in the object detection task significantly improved the performance in the cattle identification task by increasing the mAP value from 0.958 to 0.995 (Table 2). The result of this research shows that the use of data augmentation techniques can improve the performance of object detection methods.

The fact that the cattle in the dataset used in this study pass through a walkway individually rather than as a group, that all images were taken at the same time period, that all images have the same backgrounds, and that the lighting and image quality are good is advantageous for the deep learning models used in this study. It should be kept in mind that testing these models on natural images of cattle in pens or pastures will slightly decrease the identification accuracy. In general, due to these results, we propose the following strategies for training new networks;

- In studies with computer vision technologies, the number of frames between classes in the dataset should be taken into account.
- In deep learning studies such as YOLO, it should be kept in mind that it may be useful to evaluate the hyper-parameters of the models used and to use different combinations of hyper-parameters if necessary without depending on default settings.
- The presence of images with different features (lighting differences, individual or group images, distant or close images, etc.) in the data set of the model to be trained will increase the identification capacity of the model.

Conclusion

In this paper, the performance of data augmentation methods on YOLO v5s has been tested against the problem of obtaining equal image or video data for each class in image classification or object detection studies in animal husbandry. According to the results of the study, the balanced data set resulted in a ~4% increase in the mAP value and superior performance (0.995 mAP) was obtained with the YOLOv5 model trained on this data set. For future work, studies on individual identification and identification of various specific behaviors in sheep, where individual identification with computer vision technologies is more difficult than in cattle, are planned.

Acknowledgements

The authors declare no conflicts of interest.

References

Altınbilek, H. F., & Kızıl, U. (2022). Identification of Paddy Rice Diseases Using Deep Convolutional Neural Networks. *Yuzuncu Yıl University Journal of Agricultural Sciences*, *32*(4), 705-713. https://doi.org/10.29133/yyutbd.1140911

Andrew W., Greatwood C., Burghardt T. (2017). *Visual localisation and individual identification of holstein friesian cattle via deep learning*. In Proceedings of the IEEE international conference on computer vision workshops. pp. 2850-2859.

Andrew, W., Hannuna, S., Campbell, N., & Burghardt, T. (2016). Automatic individual holstein friesian cattle identification via selective local coat pattern matching in RGB-D imagery. *Proceedings - International Conference on Image Processing, ICIP*, *2016-August*, 484–488. https://doi.org/10.1109/ICIP.2016.7532404

Bati, C. T., & Ser, G. (2023). SHEEPFEARNET: Sheep fear test behaviors classification approach from video data based on optical flow and convolutional neural networks. *Computers and Electronics in Agriculture*, *204*, 107540. https://doi.org/10.1016/J.COMPAG.2022.107540

Bocaj, E., Uzunidis, D., Kasnesis, P., & Patrikakis, C. Z. (2020). On the Benefits of Deep Convolutional Neural Networks on Animal Activity Recognition. *Proceedings of 2020 International Conference on Smart Systems and Technologies, SST 2020*, 83–88. https://doi.org/10.1109/SST49455.2020.9263702

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). *YOLOv4: Optimal Speed and Accuracy of Object Detection*. https://arxiv.org/abs/2004.10934v1

Chen, Z., Wu, R., Lin, Y., Li, C., Chen, S., Yuan, Z., Chen, S., & Zou, X. (2022). Plant Disease Recognition Model Based on Improved YOLOv5. *Agronomy 2022, Vol. 12, Page 365*, *12*(2), 365. https://doi.org/10.3390/AGRONOMY12020365

Dac H.H., Gonzalez Viejo C., Lipovetzky N., Tongson E., Dunshea F. R., Fuentes S. (2022). Livestock Identification Using Deep Learning for Traceability. *Sens.* 22(21): 8256.

de Lima Weber, F., de Moraes Weber, V. A., de Moraes, P. H., Matsubara, E. T., Paiva, D. M. B., de Nadai Bonin Gomes, M., de Oliveira, L. O. F., de Medeiros, S. R., & Cagnin, M. I. (2023). Counting cattle in UAV images using convolutional neural network. *Remote Sensing Applications: Society and Environment*, *29*, 100900. https://doi.org/10.1016/J.RSASE.2022.100900

Egi, Y., Hajyzadeh, M., & Eyceyurt, E. (2022). Drone-Computer Communication Based Tomato Generative Organ Counting Model Using YOLO V5 and Deep-Sort. *Agriculture (Switzerland)*, *12*(9), 1290. https://doi.org/10.3390/AGRICULTURE12091290/S1

Huang, X., Li, X., & Hu, Z. (2019). Cow tail detection method for body condition score using Faster R-CNN. *2019 IEEE International Conference on Unmanned Systems and Artificial Intelligence, ICUSAI 2019*, 347–351. https://doi.org/10.1109/ICUSAI47366.2019.9124743

Jintasuttisak T., Edirisinghe E., Elbattay, A. (2022). Deep neural network based date palm tree detection in drone imagery. *Comput. Elect. Agricul*. 192: 106560.

Jocher G. (2020). YOLOv5. Available online: https://github.com/ultralytics/yolov5.

Kasfi, K. T., Hellicar, A., & Rahman, A. (2016). Convolutional Neural Network for time series cattle behaviour classification. *ACM International Conference Proceeding Series*, 8–12. https://doi.org/10.1145/3014340.3014342

Lee, J., Lim, K., & Cho, J. (2022). Improved Monitoring of Wildlife Invasion through Data Augmentation by Extract–Append of a Segmented Entity. *Sensors*, *22*(19). https://doi.org/10.3390/s22197383

Li, Z., Tian, X., Liu, X., Liu, Y., & Shi, X. (2022). A Two-Stage Industrial Defect Detection Framework Based on Improved-YOLOv5 and Optimized-Inception-ResnetV2 Models. *Applied Sciences 2022, Vol. 12, Page 834*, *12*(2), 834. https://doi.org/10.3390/APP12020834

Liu, S., Qi, L., Qin, H., Shi, J., & Jia, J. (2018). Path Aggregation Network for Instance Segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. https://doi.org/10.1109/CVPR.2018.00913

Luo, W., Zhang, Z., Fu, P., Wei, G., Wang, D., Li, X., Shao, Q., He, Y., Wang, H., Zhao, Z., Liu, K., Liu, Y., Zhao, Y., Zou, S., & Liu, X. (2022). Intelligent Grazing UAV Based on Airborne Depth Reasoning. *Remote Sensing 2022, Vol. 14, Page 4188*, *14*(17), 4188. https://doi.org/10.3390/RS14174188

Masebo, N. T., Marliani, G., Cavallini, D., Accorsi, P. A., Di Pietro, M., Beltrame, A., Gentile, A., & Jacinto, J. G. P. (2023). Health and welfare assessment of beef cattle during the adaptation period in a specialized commercial fattening unit. *Research in Veterinary Science*, *158*, 50–55. https://doi.org/10.1016/J.RVSC.2023.03.008

Paszke A. et al., (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Advances in Neural Information Processing Systems 32. Curran Associates, Inc. pp. 8024–8035. Available at: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

Ser, G., & Bati, C. T. (2019). Determining the best model with deep neural networks: Keras application on mushroom data. *Yuzuncu Yil University Journal of Agricultural Sciences*, *29*(3). https://doi.org/10.29133/yyutbd.505086

Shojaeipour, A., Falzon, G., Kwan, P., Hadavi, N., Cowley, F. C., & Paul, D. (2021). Automated muzzle detection and biometric identification via few-shot deep transfer learning of mixed breed cattle. *Agronomy*, *11*(11). https://doi.org/10.3390/agronomy11112365

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, *6*(1). https://doi.org/10.1186/s40537-019-0197-0

Skalski P. (2019). Make Sense. Available online: https://github.com/SkalskiP/make-sense/

Subedi, S., Bist, R., Yang, X., & Chai, L. (2023). Tracking pecking behaviors and damages of cage-free laying hens with machine vision technologies. *Computers and Electronics in Agriculture*, *204*, 107545. https://doi.org/10.1016/J.COMPAG.2022.107545

Sun, F., Wang, H., & Zhang, J. (2021). A Recognition Method of Cattle and Sheep Based on Convolutional Neural Network. *Proceedings - 2021 2nd International Seminar on Artificial Intelligence, Networking and Information Technology, AINIT 2021*, 420–424. https://doi.org/10.1109/AINIT54228.2021.00088

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual (Scotts Valley, CA: CreateSpace;).[Google Scholar]*.

Yan, B., Fan, P., Lei, X., Liu, Z., & Yang, F. (2021). A real-time apple targets detection method for picking robot based on improved YOLOv5. *Remote Sensing*, *13*(9). https://doi.org/10.3390/rs13091619

Zhang, Y., Yu, C., Liu, H., Chen, X., Lei, Y., Pang, T., & Zhang, J. (2022). An Integrated Goat Head Detection and Automatic Counting Method Based on Deep Learning. *Animals*, *12*(14). https://doi.org/10.3390/ani12141810