doi:10.20944/preprints202011.0331.v1

Article

Cattle Detection Using Oblique UAV Images

Jayme Garcia Arnal Barbedo [1]*, Luciano Vieira Koenigkan [1] and Patrícia Menezes Santos [2]

[1] Embrapa Informatica Agropecuaria; jayme.barbedo@embrapa.br;luciano.vieira@embrapa.br
[2] Embrapa Pecuaria Sudeste; patricia.santos@embrapa.br

Abstract: The evolution in imaging technologies and artificial intelligence algorithms, coupled with improvements in UAV technology, has enabled the use of unmanned aircraft in a wide range of applications. The feasibility of this kind of approach for cattle monitoring has been demonstrated by several studies, but practical use is still challenging due to the particular characteristics of this application, such as the need to track mobile targets and the extensive areas that need to be covered in most cases. The objective of this study was to investigate the feasibility of using a tilted angle to increase the area covered by each image. Deep Convolutional Neural Networks (Xception architecture) were used to generate the models for the experiments, which covered aspects like ideal input dimensions, effect of the distance between animals and sensor, effect of classification error on the overall detection process, and impact of physical obstacles on the accuracy of the model. Experimental results indicate that oblique images can be successfully used under certain conditions, but some practical limitations need to be addressed in order to make this approach appealing.

Keywords: Convolutional neural network; Unmanned Aerial Vehicles; Deep learning.

---

1. Introduction

The management of beef cattle farms operating under an extensive production system is challenging, especially considering that many of those farms have large areas with deficient communications infrastructure and ground access. Under those conditions, thorough visual inspection of the herd often requires manned flight, which are expensive and have some risks associated [1]. Because of this, horseback ground monitoring is still very common, making it very difficult to cover the entire farm in a timely manner.

Unmanned aerial vehicles (UAV) are being explored as a time- and cost-effective alternative to the approaches mentioned above. The idea is to use UAVs to capture a large number of images from a certain area, and then use algorithms to extract the information of interest. In the case of cattle monitoring, applications that have been investigated include animal detection and counting [1–5], specimen recognition [6], measurement of the distance between cow and calf [7], and determination of feeding behavior [8]. With very few exceptions [9], the information contained in the images is extracted by means of deep learning models, using one of four main approaches [1]: semantic segmentation, which associates each pixel in the image to a class; instance segmentation, which detects and delineates each distinct object of interest [4,5]; object detection, which delineates a box bounding the objects of interest [10,11]; heat mapping (probability distributions) using Convolutional Neural Networks (CNNs), which reveals the position of the animals in the image [12–14]. Different degrees of success have been achieved, a fact that is more related to the specific experimental setup and to the characteristics of the datasets used in the study than to the algorithms themselves [1]. While significant progress has been achieved, this kind of technology still cannot reach its full potential due to a number of technical and practical limitations.

One major practical limitation is that most UAVs do not have enough autonomy to cover entire farms in a single flight [15]. Depending on the size of the area to be covered, many flight missions may be required. Even if a large number of charged spare batteries is available, valuable time will be lost between flights. This is a problem because in the meantime animals may move, weather conditions

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 November 2020 doi:10.20944/preprints202011.0331.v1

Table 1. Datasets used in the experiments. The criteria used to select the samples in the "cattle" and "non-cattle" sets are described in the "Experimental Setup" section.

Size	# cattle images	# non-cattle images
224x224	276	276
112x112	856	856
56x56	1,754	1,754
28x28	3,530	3,530
14x14	8,984	8,984

may vary, and the angle of light incidence will change. All those factors have the potential to increase error rates. One possible way to counteract this problem is to capture the images at an angle. So far, all studies employing UAVs for cattle monitoring have in common the fact that images are captured perpendicularly to the ground (nadir positioning). This guarantees that all points in an image have approximately the same Ground Sample Distance (GSD), making it easier for the models to achieve their goals. Conversely, oblique images can cover much larger areas, reducing the number of flights required to cover a given area and, as a result, diminishing the harmful effects of animal movement and environmental conditions.

The use oblique images brings new challenges: the GSD varies considerably throughout the image, occlusions become more severe, the difficulty of detecting and measuring objects increases the farther they are from the camera, geometric and color distortions become more prominent in distant objects, among others. The objective of this study was to determine if and under which conditions oblique images are advantageous in comparison with those captured orthogonally. All experiments were carried out in the context of animal detection, which is an intermediate step toward animal counting. To the best of the authors' knowledge, this is the first study exploring the viability of using oblique images for cattle monitoring. Experiments were carried out using the Xception CNN architecture, which has yielded very accurate results in previous studies [1,3].

2. Material and Methods

2.1. Dataset

Images were captured at an altitude of 30 m with respect to the take-off position, using a DJI Mavic 2 Pro equipped with a 20-MPixel camera. Camera settings were all kept on automatic, except exposition, which used the presets "sunny" and "overcast" depending on weather conditions. Angles between sensor view and the orthogonal axis varied between 10° and 75°. Distances between animals and sensor varied from 30 m to more than 500 m. Animals from both Canchim and Nelore breeds were present during flights. All images were captured between 10 a.m. and 3 p.m.

The imaged areas are located at Canchim farm, São Carlos, Brazil (21°58′28″ S, 47°50′59″ W). Several different experiments involving livestock are carried out at this farm, making it possible to build a dataset representing many of the situations found in practice, ranging from unimpeded line of sights to busy environments with varying degrees of occlusion (Fig. 1). Sixty images were selected and subsequently divided into image blocks using regular grids ranging from 14-pixel to 224-pixel spacing, both horizontally and vertically. As a result, 5 datasets were generated (Table 1) and used to determined the ideal degree of granularity for animal detection in oblique images.

2.2. Experimental Setup

Blocks from the divided images were visually classified as "cattle" and "non-cattle" by an expert. A image block was classified as cattle if at least part of an animal could be unequivocally identified by the human expert. The process is thus inherently subjective, but more objective approaches were too time consuming for practical adoption. It is worth noting that blocks containing only a small portion

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 November 2020 doi:10.20944/preprints202011.0331.v1

Figure 1. Examples of image blocks (224x224 pixels) extracted from the original images captured in the field.

of an animal have little impact on the detection process (even if misidentified), which limits potential problems caused by inconsistencies associated to the visual selection process.

As expected, the number of "non-cattle" samples was much larger than the number of "cattle" samples. In order to avoid severely imbalanced classes, which can cause biased results [16], "non-cattle" samples were randomly selected from the complete set to match the number of cattle samples (Table 1).

Model training was carried out using 80% of the samples. The deep Convolutional Neural Network (CNN) used in this study was the Xception [17], which in previous studies has yielded high accuracies with a relatively small number of parameters [3]. Three different experiments were devised to investigate various aspects of the problem:

- The first experiment aimed at determining the ideal dimensions of the images to be used as inputs. Five different dimensions were tested: 224x224, 112x112, 56x56, 28x28, and 14x14 pixels. By default, the minimum image size accepted as input by the Xception model is 71x71 pixels, thus in the latter three cases the image blocks were upsampled to these dimensions. This does not alter the amount of information contained in the images, which are simply reassembled to match the input requirements. The metrics used for the assessment of detection quality were the following:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN), \quad (1)$$

$$Precision = TP/(TP + FP), \quad (2)$$

$$Recall = TP/(TP + FN), \quad (3)$$

$$F1\_Score = 2 * (Recall * Precision)/(Recall + Precision), \quad (4)$$

where TP, TN, FP and FN are the number of true positives, true negatives, false positives and false negatives, respectively. Confusion Matrices were also adopted to help visualize the results.

Table 2. Distribution of animals and image blocks containing animals in each distance class. Although the number of animals is fairly evenly distributed among classes, the number of image blocks containing animals is much larger at closer distances. Animals located more than 250 m from the sensor could not be reliably detected and were not considered.

Distance	% Total number of animals	% Total number of blocks with animals
30-50 m	37	65
50-100 m	20	18
100-250 m	43	17
over 250 m	-	-

- The second experiment aimed at determining how the distance between animals and sensor affected the detection. The distance between animals and sensor was estimated by firstly determining the length of a typical animal (in pixels) located directly below the UAV (30 meters), and then using a direct proportion to derive the distance of farther animals. This is obviously a rough estimate, given that the sizes and positions of the animals vary, but it is accurate enough for the specific purposes of this study. Samples were initially divided into four groups (Table 2): 30m - 50m, 50m - 100 m, 100m - 250m, and more than 250m, and detection accuracies were computed in each case. The latter group was later dropped because, at distances greater than 250 m, animal detection is unreliable even visually. Thus, the experiments carried out in this study only considered animals closer than 250 m from the sensor. All five image block dimensions were considered in this experiment. Only the "recall" metric was used here, because it is unfeasible to estimate distances for all "non-cattle" image blocks without a proper reference, so "precision" cannot be calculated.

- The third experiment aimed at determining the proportion of animals that are completely missed by the detection process. Animals close to the sensor usually appear in multiple image blocks (especially when block dimensions are small), while at greater distances a single image block can contain multiple animals. This implies that close animals can be successfully detected even if some misclassifications occur, while animals located farther from the sensor may be completely missed by just a single misclassification. The number of animals completely missed was determined for each block-dimension/animal-distance combination.

All experiments were carried out using a 10-fold cross-validation, with training and test sets being randomly generated in each repetition.

3. Results

Table 3 shows the results obtained using different image block sizes, and Table 4 further details recall results in terms of the distance class. Figure 1 shows the average, minimum and maximum values obtained for each cell in the confusion matrices.

Table 5 shows the proportion of animals completely missed by the detection process for each block-dimension/animal-distance pair.

4. Discussion

Global metrics such as accuracy and F1 score were somewhat similar for all image block sizes considered, with the exception of 14x14 pixels. It was observed that the latter carried too little information and is too sensitive to small spurious elements to provide reliable estimates, so unless otherwise stated, all remarks made in this section ignore the 14x14 case.

While global metrics were relatively homogeneous, "precision" and "recall" showed some clear patterns. As more focused (smaller) image blacks are used, distant animals tend to become more prominent in each sample, making the detection more likely (recall increases). On the other hand, small spurious elements also become more prominent, increasing the likelihood of misdetections (precision

 doi:10.20944/preprints202011.0331.v1

Table 3. Results obtained for each image block size. Each cell shows the maximum (top, in red), average (middle, in black) and minimum (bottom, in blue) values through the ten repetitions.

Block size	Accuracy	Precision	Recall	F1 Score
	0.87	0.92	0.87	0.87
224x224	0.85	0.85	0.85	0.85
	0.81	0.79	0.81	0.82
	0.87	0.87	0.95	0.87
112x112	0.85	0.84	0.87	0.85
	0.83	0.80	0.82	0.83
	0.85	0.84	0.90	0.86
56x56	0.84	0.82	0.88	0.85
	0.84	0.81	0.87	0.84
	0.85	0.82	0.91	0.85
28x28	0.83	0.80	0.89	0.84
	0.81	0.77	0.88	0.83
	0.71	0.70	0.78	0.71
14x14	0.67	0.65	0.76	0.70
	0.65	0.63	0.73	0.69

Table 4. Recall values for each image block and distance class.

Distance (m)	224x224	112x112	56x56	28x28	14x14
30-50	0.88	0.89	0.9	0.89	0.72
50-100	0.82	0.86	0.86	0.89	0.72
100-250	0.76	0.8	0.84	0.87	0.75

Table 5. Proportion (%) of animals completely missed for each image block and distance class.

Distance (m)	224x224	112x112	56x56	28x28	14x14
30-50	0	0	0	0	0
50-100	6	2	0	0	11
100-250	25	18	11	5	17

 doi:10.20944/preprints202011.0331.v1

224x224

	C	NC
C	85	15
NC	15	85

112x112

	C	NC
C	87	13
NC	16	84

56x56

	C	NC
C	88	12
NC	19	81

28x28

	C	NC
C	89	11
NC	22	78

14x14

	C	NC
C	76	24
NC	41	59

Figure 2. Average confusion matrices crossing actual (rows) and estimated (columns) counts, given in percentages. C and NC refer to the Cattle and Non-Cattle classes, respectively.

drops). That is exactly what was observed. This tradeoff between precision and recall explains why global metrics remained approximately the same across all block sizes.

It is also interesting to notice that recall values were not significantly affected in the case of close animals, as these are prominent even when larger image blocks are considered. For distant animals, however, recall was considerably larger when smaller image blocks were adopted. Although there were no similar experiments regarding precision values, given the almost linear tradeoff between these and recall values, it is safe to assume that precision values are directly proportional to the image block size and inversely proportional to the distance considered.

While accuracies and F1 scores were similar for all block sizes (except 14x14), the severity of the errors were not. In the case of small block sizes, false negatives ("cattle" blocks classified as "non-cattle") are not as damaging because there is a higher probability that each animal is represented by multiple blocks. This is particularly true in the case of distant animals – there was an instance in which a single missed 224x224 block contained seven animals. If the objective is to estimate the number of animals, such a misclassification would be very impactful. Indeed, the experiments indicated that the number of missed animals was considerably lower for small image blocks.

The conclusions drawn from the experimental results seem to indicate that smaller image blocks are advantageous for detecting animals located far from the sensor, as long as they carry enough information to train the model. It is important to notice, however, that architectures other than that used in the experiments (Xception) may be more or less sensitive to the many factors that can influence the results (e.g. contrast between animals and background, angle of incident light, etc.), thus it is recommended that new experiments be carried out if a different model is used. It is also worth pointing out that all animals in the images had light colors ranging from white to pale yellow or gray. Again, if other breeds are to be considered, new experiments need to be carried out.

Experimental results have shown that it is feasible to employ oblique images for detection of animals located up to 250 meters from the sensor. However, there are a few practical issues that need to be taken into consideration:

- Although occlusions are a problem in orthogonal images, in the case of oblique images obstacles such as trees, tall grass and sheds can obstruct a considerably larger portion of the field of view. The actual percentage of the area that can be properly scanned can vary considerably depending on the type of vegetation and where potential obstacles are located with respect to the camera. The configuration of the vegetation canopies also plays a major role, as in some cases it is better to keep major obstacles close to the camera, so the area occluded is roughly the same as the area of the canopy itself, while in other cases a better field of view can be obtained if those obstacles are farther from the camera (e.g. if branches and leaves are high above ground). In the case of the study, most of the obstacles were trees typical from the Cerrado region, which tend to have low leaf density, although there were also a few trees with dense canopies (Fig. 1). Because obstacle characteristics can vary wildly from site to site, in most cases the ideal positioning for the sensor will be specific to each area. This problem is much less prominent in farms with large obstructed areas, making it easier to explore the benefits of oblique images.

- With mobile targets, the objects of interest may not be at the same positions when new images are captured. This is true independently of the type of image used, but in the case of oblique images there is the additional challenge of trying to determine the borders of the region covered in a given image in order to avoid overlap when a new image is considered. In areas with high cattle population density, it may be nearly impossible to determine which animals have already been counted, especially if those are concentrated near the 250 m limit of detection.

- Separating clustered animals is arguably the most challenging task when counting animals using orthogonal images [1,3]. With oblique images, the problem becomes even more difficult. With tilted angles, clustering tends to intensify and the degree of overlap between animals tends to grow sharply. The problem becomes even more difficult considering that distant animals are depicted by fewer pixels, causing animal shapes to become less discernible. As a result, cluster separation may quickly become unfeasible, even visually. It is worth pointing out that the potential impact of clustered animals on the accuracy of a counting estimate was not investigated in this study, but it was thoroughly investigated in [1]. Many of the conclusions of that study hold here, but since the separation of animals far from the sensor probably pose a significantly more difficult challenge, this is an issue that should be investigated in depth in the near future.

The practical difficulties mentioned above can severely limit the appeal of using images captured at a tilted angles for cattle monitoring. It is worth pointing out, however, that the applicability of this approach can be greatly improved by simple measures such as including some markers that can be easily identified in the images, helping to delimit the region considered in each image. With this, it may become possible to devise strategies to minimize other problems, such as decreasing the distance at which successive images are captured and selecting positions that minimize the impact of obstacles. Another action that can improve the results obtained by this kind of approach is to explore the prior knowledge about topography and obstacles to determine the best spots for capturing the images. These are issues worth exploring in future studies, together with an investigation to determine if the potential benefits brought by oblique images are enough to justify the additional effort needed to enable their use.

5. Conclusion

This article explored the possibility of using tilted angles to increase the area covered by a single image captured using UAVs. Experimental results indicated that this approach can be advantageous if challenges related to view obstructions and to the determination of the exact borders of the region considered in the image can be properly addressed. Future investigations should include a cost-benefit analysis to estimate the potential benefits of oblique images against the measures needed to minimize practical hurdles.

This study dealt only with the problem of animal detection, which is only the first step of more complex tasks, such as animal counting and detection of anomalies. These have several technical challenges (separation of clustered animals, animal tracking, etc. [1,3]) which are exacerbated when dealing with oblique images. Future research should also tackle these challenges in order to enable technologies capable of providing the answers needed by decision makers.

6. Acknowledgements

This work was funded by Fapesp under grant number 2018/12845-9. The authors would also like to thank Nvidia for donating the GPU used in the experiments.

1. Barbedo, J.G.A.; Koenigkan, L.V.; Santos, P.M.; Ribeiro, A.R.B. Counting Cattle in UAV Images—Dealing with Clustered Animals and Animal/Background Contrast Changes. *Sensors* **2020**, *20*. doi:10.3390/s20072126.
2. Goolsby, J.A.; Jung, J.; Landivar, J.; McCutcheon, W.; Lacewell, R.; Duhaime, R.; Baca, D.; Puhger, R.; Hasel, H.; Varner, K.; Miller, B.; Schwartz, A.; Perez de Leon, A. Evaluation of Unmanned Aerial Vehicles (UAVs) for detection of cattle in the Cattle Fever Tick Permanent Quarantine Zone. *Subtropical Agriculture and Environments* **2016**, *67*, 24–27.
3. 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*. doi:10.3390/s19245436.
4. Xu, B.; Wang, W.; Falzon, G.; Kwan, P.; Guo, L.; Chen, G.; Tait, A.; Schneider, D. Automated cattle counting using Mask R-CNN in quadcopter vision system. *Computers and Electronics in Agriculture* **2020**, *171*, 105300. doi:https://doi.org/10.1016/j.compag.2020.105300.
5. Xu, B.; Wang, W.; Falzon, G.; Kwan, P.; Guo, L.; Sun, Z.; Li, C. Livestock classification and counting in quadcopter aerial images using Mask R-CNN. *International Journal of Remote Sensing* **2020**, *TBD*, 1–22. doi:10.1080/01431161.2020.1734245.
6. Andrew, W.; Gao, J.; Mullan, S.; Campbell, N.; Dowsey, A.W.; Burghardt, T. Visual Identification of Individual Holstein-Friesian Cattle via Deep Metric Learning. *ArXiv* **2020**, *abs/2006.09205*.
7. Mufford, J.T.; Hill, D.J.; Flood, N.J.; Church, J.S. Use of unmanned aerial vehicles (UAVs) and photogrammetric image analysis to quantify spatial proximity in beef cattle. *Journal of Unmanned Vehicle Systems* **2019**, *7*, 194–206. doi:10.1139/juvs-2018-0025.
8. Nyamuryekung'e, S.; Cibils, A.; Estell, R.; Gonzalez, A. Use of an Unmanned Aerial Vehicle – Mounted Video Camera to Assess Feeding Behavior of Raramuri Criollo Cows. *Rangeland Ecology & Management* **2016**, *69*, 386–389.
9. Longmore, S.; Collins, R.; Pfeifer, S.; Fox, S.; Mulero-Pázmány, M.; Bezombes, F.; Goodwin, A.; Juan Ovelar, M.; Knapen, J.; Wich, S. Adapting astronomical source detection software to help detect animals in thermal images obtained by unmanned aerial systems. *International Journal of Remote Sensing* **2017**, *38*, 2623–2638.
10. Andrew, W.; Greatwood, C.; Burghardt, T. Aerial Animal Biometrics: Individual Friesian Cattle Recovery and Visual Identification via an Autonomous UAV with Onboard Deep Inference. *ArXiv* **2019**, *abs/1907.05310v1*.
11. Shao, W.; Kawakami, R.; Yoshihashi, R.; You, S.; Kawase, H.; Naemura, T. Cattle detection and counting in UAV images based on convolutional neural networks. *International Journal of Remote Sensing* **2020**, *41*, 31–52.
12. Chamoso, P.; Raveane, W.; Parra, V.; González, A. UAVs Applied to the Counting and Monitoring of Animals. Advances in Intelligent Systems and Computing. *Advances in Intelligent Systems and Computing* **2014**, *291*, 71–80.
13. Rahnemoonfar, M.; Dobbs, D.; Yari, M.; Starek, M. DisCountNet: Discriminating and Counting Network for Real-Time Counting and Localization of Sparse Objects in High-Resolution UAV Imagery. *Remote Sensing* **2019**, *11*, 1128.
14. Rivas, A.; Chamoso, P.; González-Briones, A.; Corchado, J. Detection of Cattle Using Drones and Convolutional Neural Networks. *Sensors* **2018**, *18*, 2048.
15. Barbedo, J.G.A.; Koenigkan, L.V. Perspectives on the use of unmanned aerial systems to monitor cattle. *Outlook on Agriculture* **2018**, *47*, 214–222. doi:10.1177/0030727018781876.
16. Buda, M.; Maki, A.; Mazurowski, M.A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks* **2018**, *106*, 249–259. doi:https://doi.org/10.1016/j.neunet.2018.07.011.
17. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. *ArXiv* **2017**, *abs/1610.02357v3*.