
Counting Cows: Tracking Illegal Cattle Ranching From High-Resolution Satellite Imagery

Issam Laradji^{1,2,5}, Pau Rodriguez², Freddie Kalaitzis³, David Vazquez², Ross Young², Ed Davey⁴, and Alexandre Lacoste²

¹issam.laradji@gmail.com, ²Element AI, ³Oxford / FDL, ⁴Global Witness, ⁵McGill University

Abstract

Cattle farming is responsible for 8.8% of greenhouse gas emissions worldwide. In addition to the methane emitted due to their digestive process, the growing need for grazing areas is an important driver of deforestation. While some regulations are in place for preserving the Amazon against deforestation, these are being flouted in various ways, hence the need to scale and automate the monitoring of cattle ranching activities. Through a partnership with *Global Witness*, we explore the feasibility of tracking and counting cattle at the continental scale from satellite imagery. With a license from Maxar Technologies, we obtained satellite imagery of the Amazon at 40cm resolution, and compiled a dataset of 903 images containing a total of 28498 cattle. Our experiments show promising results and highlight important directions for the next steps on both counting algorithms and the data collection process for solving such challenges. The code is available at https://github.com/IssamLaradji/cownter_strike.

1 Introduction

With a population nearing 1 billion, cattle are the 2nd most populous mammal on earth, after humans. Cattle farming produces 8.8% of greenhouse gas emissions worldwide [8]¹. Besides the methane released due to digestion process of cattle, the industry demands massive regions to be deforested to make space for the grazing areas. Deforestation releases most carbon stored in the forest, it displaces indigenous communities and species, and the resulting lack of biodiversity threatens to disrupt the ecosystem. The lack of rainforests compounds the effects of global warming, and further accelerates the ongoing 6th mass extinction of species [4, 5, 21, 29].

One of the most important areas affected is the Brazilian Amazon. It is estimated that 70% of its cleared forests are now populated by cattle, ranking Brazil's herd as the 2nd largest in the world. Despite strict laws in Brazil against the rearing and selling of cattle in protected areas of the Amazon, these are widely flouted. A common workaround to the current legislation is through smuggling the cattle to the southern grasslands, where they are declared and slaughtered in a legal area. The cattle laundering network shields cattle rearing from the law even in protected regions of the Amazon, which has been further diminished by wildfires triggered to expand the grazing areas.

This trend calls for an automated scalable way to monitor against illegal cattle ranching, a task that we show to be feasible through high-res satellite images and recent advances in deep learning. Adult cattle span 2 meters, so to achieve a reliable detection we require satellite imagery at sub-meter resolution. This could be achieved with drone imagery, but with a limited range they would be too

¹In its 2013 report, the Food and Agriculture Organization [8] estimates that the livestock industry accounts for 14.5% of anthropogenic green house gas emissions, to which, beef and dairy production contribute 40% and 21% respectively.

expensive to operate on a continental-scale. On the other hand, modern satellites can be tasked for multiple, ad-hoc revisits and their on-board cameras can now resolve objects at the sub-meter level (for example, WorldView3 and SkySat constellations). In this paper we use WorldView3 (Maxar) imagery of Amazon ranches, with panchromatic (PAN) and multi-spectral (MUL) bands at 0.31 and 1.24 m/pixel respectively. Typically, MUL bands are already *pan-sharpened* to the panchromatic resolution. The acquisitions from WorldView3 are task-based, so coverage can be sparse temporally and spatially, but since August 2014 most areas of the rainforest were covered a few times.

2 Related Work

Our work touches on several fronts - remote sensing applied to animal conservation, localization and counting methods in ML, and environmental / human rights issues driven by deforestation.

Remote sensing of cattle and other animals Most of the work on the tracking and counting of animals and other small objects relies mostly on UAV/drone images [22, 31]. On cattle detection, CNNs trained with UAV images can be effective in detection [23, 2] and counting [31, 26, 3]. Animals of similar size, like wildebeest, can be counted from UAV images [27], and even animals as small as sheep [30]. Satellite images have been broadly used for animal conservation. Namely, WorldView2 panchromatic 50cm imagery was used to count yaks [32, 30]. However, satellite cameras hit a limit when it comes to resolving animals as small as a sheep [30]. We are the first to demonstrate cattle counting from space, which contributes to the much needed evidence on the animal counting capacity of cutting-edge 0.3m sensors, made available commercially just in 2014.

Localization and counting methods *Object counting* is the task of estimating the number of objects of interest in an image. This might also include localization where the goal is to identify the locations of the objects in the image. This task has important real-life applications such as ecological surveys [1, 24, 16] and cell counting [7, 12, 11]. The datasets used for counting tasks [9, 1] are often labeled with point-level annotations where a single pixel is labeled for each object. There are three main approaches to object counting: regression, density-based and detection-based methods. Regression-based methods, such as Glance [6], explicitly learn to count using image-level labels only. Density-based methods [17, 18] transform the point-level annotations to a density map through a Gaussian kernel. They are then trained to predict the density maps through with a least-squares loss. Detection-based methods [13, 15, 14] first pinpoint the objects' locations in the images and then count the number of detected instances. LCFCN [13] uses a fully convolutional neural network (FCN) and a detection-based loss function that encourages the model to output a single blob per object. In this work we evaluate a density based method, CSRNet [18], and a detection method, LCFCN [13], in the task of cattle counting and localization from satellite images.

3 Dataset

Image collection process While some satellite constellations offer global coverage with periodic revisits, Maxar's coverage is tasked-based, that is, covered regions are chosen on customer demand, and revisits occur sparsely and irregularly. As such, regions with few revisits since 2008 are commonplace. The quality of these images is far from uniform. These are compounded, for instance, by cloud coverage which varies depending on the season. The image resolution can vary, either due to an off-nadir angle of acquisition (nadir = vertical), or due to sourcing from an older sensor (e.g. WorldView2). To ensure a high overall quality, we use images with 40cm or better resolution (after pansharpening), and with less than 20% cloud coverage. To increase the chances of observing cattle, we geo-referenced the top selling ranches in Brazil² and queried Maxar's catalog around these locations.

Labelling The regions collected from Maxar's catalog were sliced into patches of 500×500 pixels. Using a custom-made labelling tool, cattle are located with point-annotations. Out of the 12,252 labelled patches, only 903 contain cattle, and the rest are labelled as "no cow". The distribution of the counts is shown in Figure 1 (right panel). Interestingly, many of the 500×500 patches contain

²Source of vector data: [Brazilian Rural Environmental Register](#)



Figure 1: Peak into the dataset. **Left:** Typical image with cattle. Image ©2020 Maxar Technologies. **Center left:** Crowded ranch. Image ©2020 Maxar Technologies. **Center right:** There is likely a white cow in the middle, while the rest of the white blobs appear due to lack of vegetation. Image ©2020 Maxar Technologies. **Right:** Histogram of the number of cattle per image.

over 100 head of cattle, and some contain over 1000 head of cattle. In total, we have labelled 28498 head of cattle.

Dataset License We plan to release this dataset for reproducibility and further machine learning development. Details of the license are still under discussion with Maxar Technologies.

4 Cattle counting with deep learning

We localize and count cattle in satellite images with two different approaches:

CSRNet [18] is a density-based method for counting and localization that requires point-level annotations for training. Density methods transform the point-level annotations into a density map through a Gaussian kernel. Then they are trained to predict the density maps, with a least-squares objective. Density-based methods often assume a fixed object size (defined by the Gaussian kernel), often in a controlled environment, which makes them difficult to use on objects of different shapes and sizes. These methods perform well for counting large amounts of objects in crowded situations.

LCFCN [13] is a detection-based method for counting and localization (LC) which can be trained with point-level annotations. It uses a fully convolutional neural network (FCN) and a detection-based loss function that encourages the model to output a single blob per object. During the training phase, the model learns to split the blobs that contain more than one point annotation and to remove the blobs that contain no point-level annotations. This method provides an accurate localization of the objects, it is robust to differences in scale, and performs well in non-crowded situations.

Our methods use a VGG16 FCN8 network [20] pretrained on ImageNet. The models are trained with a batch size of 8 for 100 epochs with ADAM [10] and a learning rate of 10^{-4} . We also achieve similar results with optimizers that do not require a learning rate [28, 19]. We use early-stopping on the validation set and report the scores on the test set.

5 Experiments

In this section we assess the viability of locating and counting cattle from satellite images, as well as the weaknesses of current approaches. We compare between ImageNet pretrained-CSRNet and -LCFCN and an LCFCN with weights initialized using the Xavier method on the following two different metrics:

Mean Absolute Percentage Error (MAPE) Given ground-truth counts y and estimated counts \hat{y} ,

$$\text{MAPE} = \frac{1}{n} \sum_i |y_i - \hat{y}_i| / \max(y_i, 1), \quad (1)$$

where index i runs over the evaluation set, and \max prevents division by zero ground-truth counts. This variant of the mean absolute error (MAE) allows the comparison of crowded vs. sparse regions.

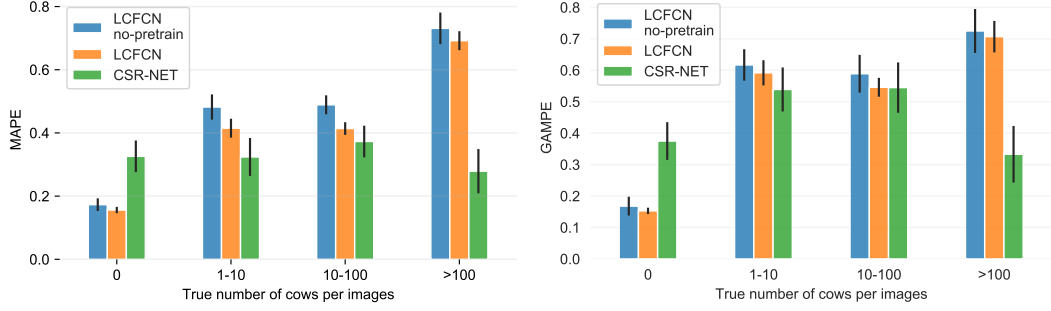


Figure 2: Counting and localization results for different cattle densities (lower is better). Error bars are ± 1 std. dev. over 3 runs of the learning algorithm with different random seeds. **Left:** Mean Absolute Percentage Error (MAPE). **Right:** Grid Average Mean absolute Percentage Error (GAMPE).

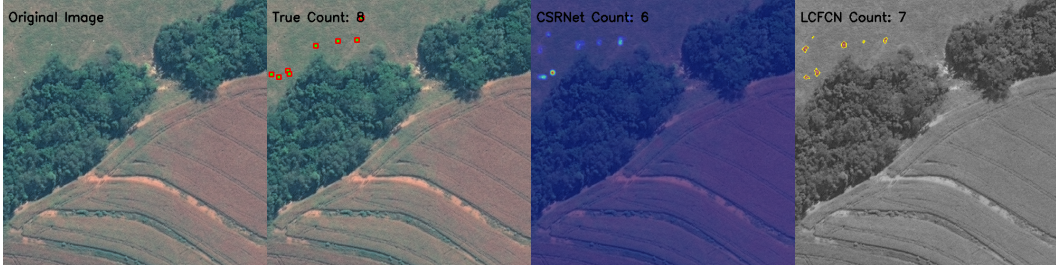


Figure 3: Typical examples of results. **Left:** Original, **Center left:** Ground-truth annotation. **Center right:** CSRNet results. **Right:** LCFCN results. Image ©2020 Maxar Technologies.

Grid Average Mean absolute Percentage Error (GAMPE) measures counting and localization performance. GAMPE is a variant of GAME [9] which uses MAPE instead of MAE. GAMPE partitions the image using a grid of non-overlapping sub-regions, and the error is computed as the sum of the mean absolute percentage errors in each of these sub-regions. This metric tends to penalize methods that rely on spurious correlations to guess the count, instead of relying on informative features.

Figure 2 reports the MAPE and GAMPE on images with different cattle densities. The left-most bin (0 cattle per image) is made up of images with no cattle. In this case LCFCN outperforms CSRNet. In a binary (cattle vs no-cattle) classification benchmark, LCFCN achieves an F-score of 0.676 whereas CSRNet achieved 0.571. This result suggests that LCFCN is best suited at assessing the presence vs. absence of cattle. We also experiment with a standard ResNet-50 classifier pretrained on ImageNet, which achieves a poor F-score of 0.210 compared to LCFCN and CSRNet.

For the rest of the bins, where there are cattle present in the images (1-10 cattle per image, or more), CSRNet consistently outperforms LCFCN both on counting and localization. Note that the performance gap of CSRNet over LCFCN grows for the densest images (>100 head of cattle). We also see that LCFCN pretrained on ImageNet tends to outperform the non-pretrained LCFCN. Figure 3 shows a typical image from the dataset with the annotations and the results from the proposed methods.

6 Conclusion

We showed that cattle detection from modern high-resolution satellite sensors is achievable, but further work is needed for a large-scale deployment. Our results suggest that a hybrid model of LCFCN and CSR-NET has merit. However, we believe that the biggest improvement would come from incorporating more information. Specifically, if we use different revisits of the same geo-location, it would be easier to distinguish cattle from static objects that resemble cattle, like bushes, rocks, and patches of sand. This would enhance the labelling quality and the counting algorithm.

Broader Impact

The same methodology can be applied for tracking any large animals that span more than 1 meter from the sky. This could be valuable for the census of wildlife. Ecologists could monitor animal populations more frequently, at lower costs, and act faster in the face of a species extinction.

On the flip side, high-resolution satellite imagery can be abused for surveillance[25]. We believe that the main limiting factor to such applications is the access to high-resolution images, and not the development of counting algorithms. Also, the identification of a specific individual requires resolutions that are not yet available in commercial satellites.

References

- [1] C. Arteta, V. Lempitsky, and A. Zisserman. Counting in the wild. *ECCV*, 2016.
- [2] J. G. A. Barbedo, L. V. Koenigkan, T. T. Santos, and P. M. Santos. A study on the detection of cattle in uav images using deep learning. *Sensors*, 19(24):5436, 2019.
- [3] J. G. A. Barbedo, L. V. Koenigkan, P. M. Santos, and A. R. B. Ribeiro. Counting cattle in uav images—dealing with clustered animals and animal/background contrast changes. *Sensors*, 20(7):2126, 2020.
- [4] B. W. Brook, N. S. Sodhi, and P. K. Ng. Catastrophic extinctions follow deforestation in singapore. *Nature*, 424(6947):420–423, 2003.
- [5] B. W. Brook, C. J. Bradshaw, L. P. Koh, and N. S. Sodhi. Momentum drives the crash: mass extinction in the tropics. *Biotropica*, 38(3):302–305, 2006.
- [6] P. Chattopadhyay, R. Vedantam, R. RS, D. Batra, and D. Parikh. Counting everyday objects in everyday scenes. *CVPR*, 2017.
- [7] J. P. Cohen, G. Boucher, C. A. Glastonbury, H. Z. Lo, and Y. Bengio. Count-ception: Counting by Fully Convolutional Redundant Counting. *ICCV Workshops*, 2017.
- [8] P. J. Gerber, H. Steinfeld, B. Henderson, A. Mottet, C. Opio, J. Dijkman, A. Falcucci, G. Tempio, et al. *Tackling climate change through livestock: a global assessment of emissions and mitigation opportunities*. Food and Agriculture Organization of the United Nations (FAO), 2013.
- [9] R. Guerrero, B. Torre, R. Lopez, S. Maldonado, and D. Onoro. Extremely overlapping vehicle counting. *IbPRIA*, 2015.
- [10] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [11] I. Laradji, P. Rodriguez, F. Branchaud-Charron, K. Lensink, P. Atighehchian, W. Parker, D. Vazquez, and D. Nowrouzezahrai. A weakly supervised region-based active learning method for covid-19 segmentation in ct images. *arXiv preprint arXiv:2007.07012*, 2020.
- [12] I. Laradji, P. Rodriguez, O. Manas, K. Lensink, M. Law, L. Kurzman, W. Parker, D. Vazquez, and D. Nowrouzezahrai. A weakly supervised consistency-based learning method for covid-19 segmentation in ct images. In *Winter Applications in Computer Vision (WACV)*, 2021.
- [13] I. H. Laradji, N. Rostamzadeh, P. O. Pinheiro, D. Vazquez, and M. Schmidt. Where are the blobs: Counting by localization with point supervision. *ECCV*, 2018.
- [14] I. H. Laradji, D. Vazquez, and M. Schmidt. Where are the masks: Instance segmentation with image-level supervision. In *British Machine Computer Vision (BMVC)*, 2019.
- [15] I. H. Laradji, R. Pardinas, P. Rodriguez, and D. Vazquez. Looc: Localize overlapping objects with count supervision. In *International Conference on Image Processing (ICIP)*, 2020.
- [16] I. H. Laradji, A. Saleh, , P. Rodriguez, M. Nowrouzezahrai, Derek Rahimiazghadi, and D. Vazquez. Affinity lfcn: Learning to segment fish with weak supervision. *Arxiv preprint arXiv:*, 2020.
- [17] V. Lempitsky and A. Zisserman. Learning to count objects in images. *NIPS*, 2010.
- [18] Y. Li, X. Zhang, and D. Chen. Csrnet: Dilated convolutional neural networks for understanding the highly congested scenes. *CVPR*, 2018.

- [19] N. Loizou, S. Vaswani, I. Laradji, and S. Lacoste-Julien. Stochastic polyak step-size for sgd: An adaptive learning rate for fast convergence. *arXiv preprint arXiv:2002.10542*, 2020.
- [20] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. *CVPR*, 2015.
- [21] J. A. Pounds, M. R. Bustamante, L. A. Coloma, J. A. Consuegra, M. P. Fogden, P. N. Foster, E. La Marca, K. L. Masters, A. Merino-Viteri, R. Puschendorf, et al. Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature*, 439(7073):161–167, 2006.
- [22] M. Rahnemounfar, D. Dobbs, M. Yari, and M. J. Starek. Discountnet: Discriminating and counting network for real-time counting and localization of sparse objects in high-resolution uav imagery. *Remote Sensing*, 11(9):1128, 2019.
- [23] A. Rivas, P. Chamoso, A. González-Briones, and J. M. Corchado. Detection of cattle using drones and convolutional neural networks. *Sensors*, 18(7):2048, 2018.
- [24] A. Saleh, I. H. Laradji, D. A. Konovalov, M. Bradley, D. Vazquez, and M. Sheaves. A realistic fish-habitat dataset to evaluate algorithms for underwater visual analysis. *Scientific Reports*, 10(1):1–10, 2020.
- [25] C. Santos and L. Rapp. Satellite imagery, very high-resolution and processing-intensive image analysis: Potential risks under the gdpr. *Air and Space Law*, 44:275–296, 2019.
- [26] W. Shao, R. Kawakami, R. Yoshihashi, S. You, H. Kawase, and T. Naemura. Cattle detection and counting in uav images based on convolutional neural networks. *International Journal of Remote Sensing*, 41(1): 31–52, 2020.
- [27] C. J. Torney, A. P. Dobson, F. Borner, D. J. Lloyd-Jones, D. Moyer, H. T. Maliti, M. Mwita, H. Fredrick, M. Borner, and J. G. C. Hopcraft. Assessing rotation-invariant feature classification for automated wildebeest population counts. *Plos one*, 11(5):e0156342, 2016.
- [28] S. Vaswani, A. Mishkin, I. Laradji, M. Schmidt, G. Gidel, and S. Lacoste-Julien. Painless stochastic gradient: Interpolation, line-search, and convergence rates. In *Advances in Neural Information Processing Systems*, pages 3732–3745, 2019.
- [29] P. M. Vitousek. Beyond global warming: ecology and global change. *Ecology*, 75(7):1861–1876, 1994.
- [30] D. Wang, Q. Song, X. Liao, H. Ye, Q. Shao, J. Fan, N. Cong, X. Xin, H. Yue, and H. Zhang. Integrating satellite and unmanned aircraft system (uas) imagery to model livestock population dynamics in the longbao wetland national nature reserve, china. *Science of The Total Environment*, 746:140327, 2020.
- [31] B. G. Weinstein. A computer vision for animal ecology. *Journal of Animal Ecology*, 87(3):533–545, 2018.
- [32] Y. Xue, T. Wang, and A. K. Skidmore. Automatic counting of large mammals from very high resolution panchromatic satellite imagery. *Remote sensing*, 9(9):878, 2017.