

Man-made Structures Detection from Space

Fundamentals of Data Science Master's Thesis

Eduard Ribas, Peter Weber

Advisor: Dr. Jordi Vitrià¹

¹Departament de Matemàtiques i Informàtica,
Universitat de Barcelona

Barcelona, July 2, 2019



Initial frame

Just some initial content...



Outline

1 Introduction

- Motivation
- Satellogic
- Thesis goal and outline

2 Datasets

- USGS datasource

3 Deep Learning

- ResNet

4 Proposed approach

- ResNet activations

5 Results

- Man-made structures detection at different scales

6 Conclusions

- The Problem
- The Datasets
- The Models and Results
- Future work

References



1 Introduction

2 Datasets

3 Deep Learning

4 Proposed approach

5 Results

6 Conclusions

References



Motivation I

Human kind exerts an ever increasing pressure on natural and ecological systems due to the associated consequences of the explosion of human population. The exploitation of the earth manifests itself in extraction of natural resources, proliferation of human-made infrastructure and waste, and increasing production land use for crop and pasture land [1]. As a logical consequence, we observe widespread declines in biodiversity [2], decrease in natural habitat, attrition of wilderness areas, deforestation, and enhanced emission of greenhouse gases to the atmosphere. This increasing intrusion leads to reduction of benefits that humans receive from natural systems [3] such as the extinction of natural resources, and ultimately provoke natural disaster induced by effects such as climate change.



This work has been developed in cooperation with Satellogic, a company that provides earth observation data and analytics as a service to enable better decision making for industries, governments, and individuals. Satellogic was founded in 2010 in Buenos Aires, and has expanded since then with offices in Barcelona and China. Satellogic builds, launches and maintains their own satellites.



Thesis goal and outline I

The motivation for this Master's thesis is to provide an answer to the question:

What is the optimal resolution to detect human impact in satellite imagery, having in mind the economical cost of acquiring and processing the information?



1 Introduction

2 Datasets

3 Deep Learning

4 Proposed approach

5 Results

6 Conclusions

References



USGS datasource I

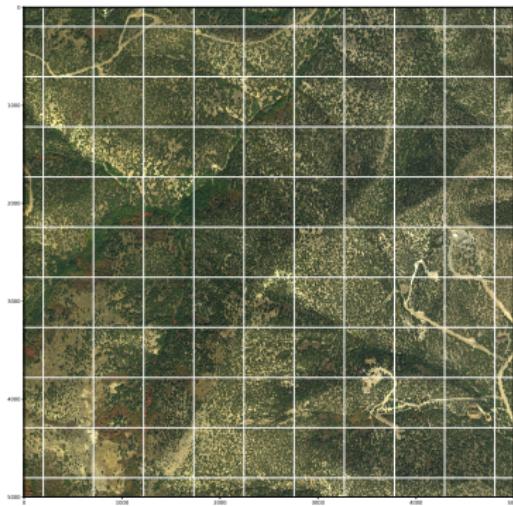


Figure: Example of unprocessed image. This image has a size of 5000×5000 pixels. The continuous white lines show how we crop smaller images of size 512×512 pixels from the original one.

1 Introduction

2 Datasets

3 Deep Learning

4 Proposed approach

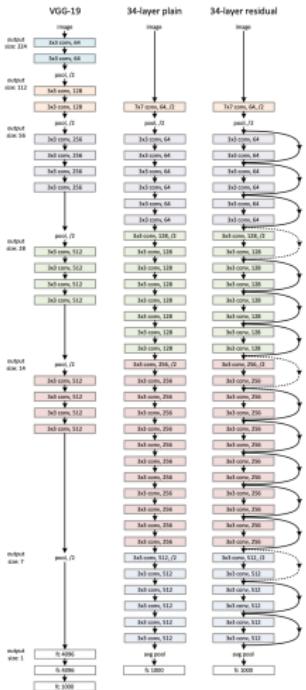
5 Results

6 Conclusions

References



ResNet



ResNet architecture

Figure: Comparison between
deep Convolutional Neural
Networks



1 Introduction

2 Datasets

3 Deep Learning

4 Proposed approach

5 Results

6 Conclusions

References



ResNet activations I

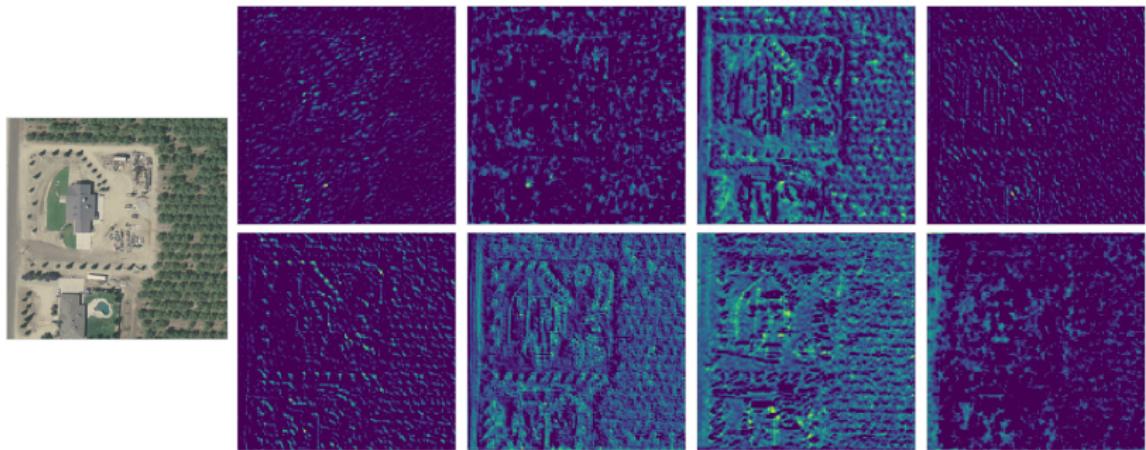


Figure: ResNet activations of an Agriculture image: 10th layer.



ResNet activations II

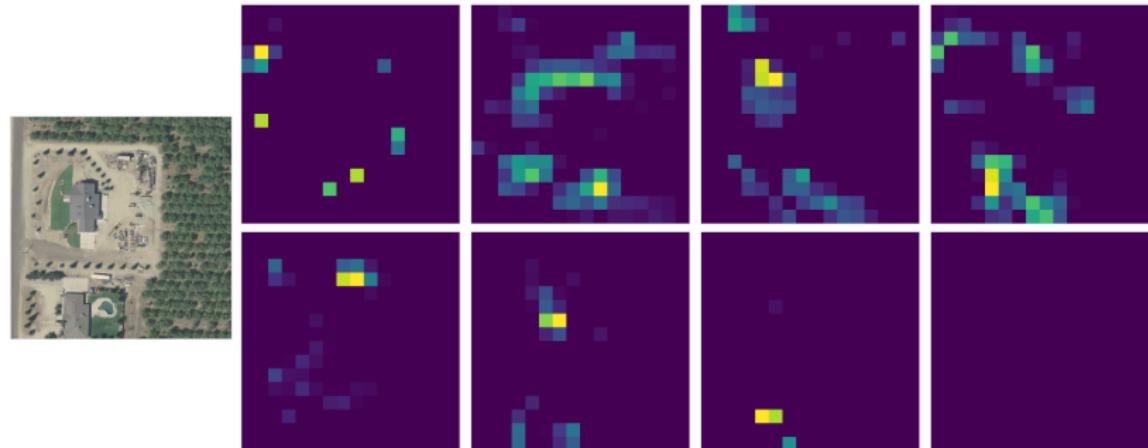
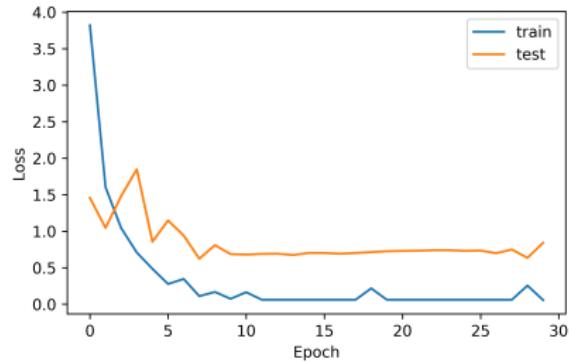
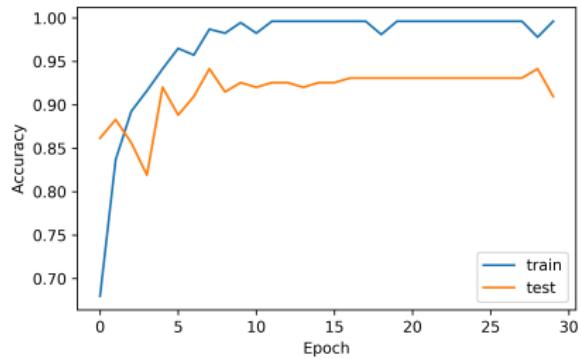


Figure: ResNet activations of an Agriculture image: final layer, 49th.

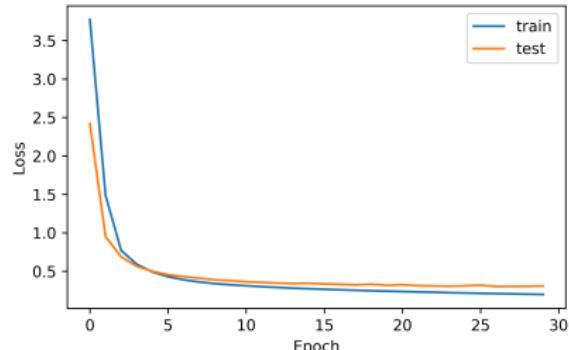
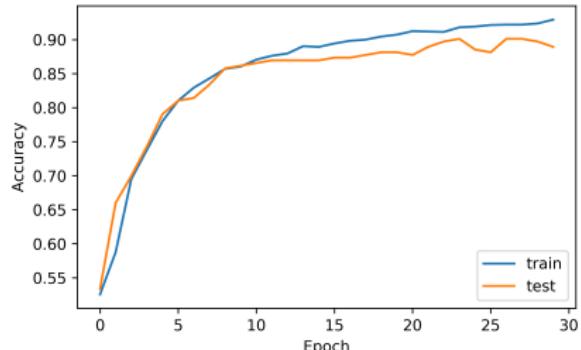


ResNet activations III

Dataset: 0.3m, Resolution: 0.3m, Neurons: 100



Dataset: 0.3m, Resolution: 4.8m, Neurons: 100



1 Introduction

2 Datasets

3 Deep Learning

4 Proposed approach

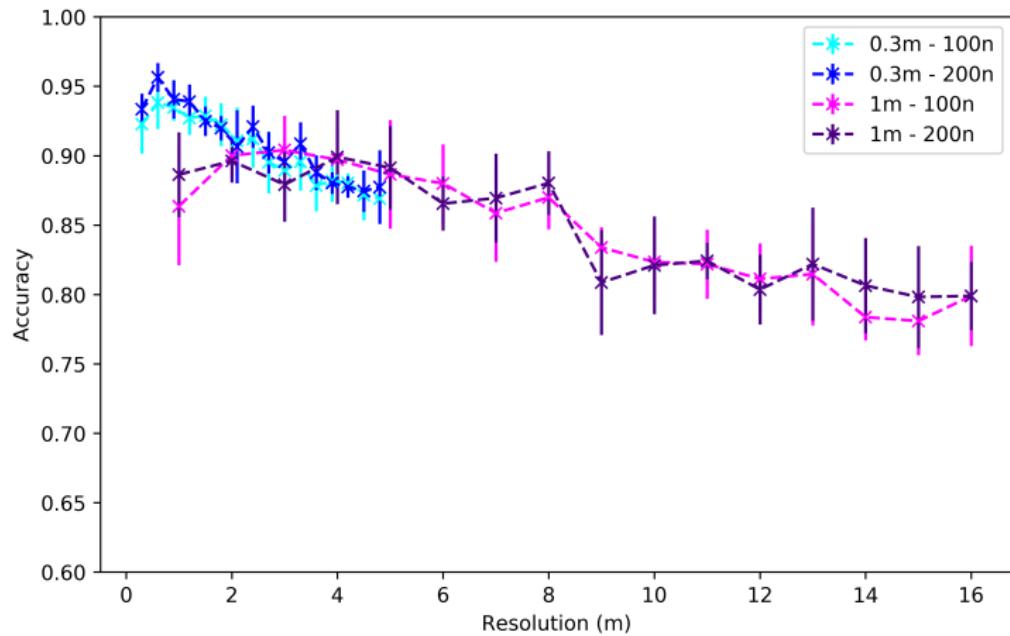
5 Results

6 Conclusions

References



Man-made structures detection at different scales I



Man-made structures detection at different scales II

Figure: Accuracy at each resolution for all datasets and architectures.

Similar results are obtained at the same degraded resolutions. The vertical lines represent the variability (standard deviation) for the different cross-validation folds.



1 Introduction

2 Datasets

3 Deep Learning

4 Proposed approach

5 Results

6 Conclusions

References

The Problem I

The initial phase during the development of the project consisted on clearly defining the problem to study. The goal was to investigate which satellite imagery resolutions allowed for an accurate detection of man-made structures, and what would be the cost associated. For that, we needed to define the scope of human activity to consider, look for suitable datasets for this study (which eventually lead to building our own datasets), and define the actual technical problem to be modeled, in order to evaluate the accuracy by resolution.

Both for the datasets and the problem, we needed them to be feasible enough to not require high technical and computational efforts (which would be, for instance, trying to detect every type of human activity in the images, providing their position and shape, and classifying them into several more categories). On the other hand, we needed it to be a realistic situation, so the results obtained could be extrapolated to other, more complex scenarios.



The Problem II

All in all, having a well-defined problem scope and a good approach to tackle it allowed us to achieve remarkable and realistic results.



The Datasets I

After investigating existing datasets of labeled satellite images, we could not find the one suitable for our purpose, as most of them were mainly focused towards urban areas, or were just build for some other different goal that would not work for us. Hence, we decided to build it ourselves. It had to be representative enough to pose a challenge for our models, yet feasible to be build with our available time and resources. The four categories considered (agriculture, shrubland-grassland, semi-desert, and forest-woodland) and the balance between non-existent and existent human impact images allowed us to build a good and representative dataset, which makes reasonable to eventually generalize the results obtained to other scenarios.

Of course, we acknowledge that having a larger dataset of images, annotated with higher degree of detail, like position, shapes and types of man-made or natural structures, would be great to build high performance models, capable of detecting all sorts of human impact with far better



The Datasets II

accuracy. Nevertheless, this high-precision goal could not be fitted into our general purpose.



The Models and Results I

Using a pre-trained CNN like the ResNet helped us to speed up the training process and achieve good results without requiring a very large dataset and computational power. The binary classification problem considered turned out to be feasible and representative of how accuracy is affected with a decrease in the image resolution.

From the results we observe that, as expected, the higher the resolution the better, but also that there seems to be sweet spot between $1-2m$ and $8m$ where, except for the more subtle forms of human activity, most of the man-made structures studied are detectable with good accuracy. This trade-off with the resolution allows to consider more cost-economic satellite solutions without dramatically compromising accuracy and utility. For instance, operating a satellite at $2m$ (or $8m$) instead of $1m$ reduces the cost approximately by a factor 6 (or 100).

Future work I

With all that said, we realize that there is still plenty of space for further work and investigations, so let us now indicate some of these ideas.

- First of all, having a better dataset could help improving the results and open new investigation lines to explore. It could be improved and enlarged with more variate images, with a better and more consistent classification, and including more detailed annotations of the position and type of objects or structures appearing. This would allow to train more accurate models capable of detecting all these kind of human impact.



Future work II

- Regarding the model, other techniques for feature extraction could be studied, like other pre-trained Neural Networks, and the parameters itself (like the number of activations to consider, the architecture of the model or the training phase) could be further fine-tuned. Additionally, the pipeline could be made more robust so that it could ingest a larger amount of data, as part of the improvements suggested for the dataset. And, of course, having a powerful computational cluster would allow to speed up the processes and target more ambitious goals.
- A more in depth study of the results could help to understand more precisely on which images the algorithms fail, what kind of information are learning (patterns, colors, shapes, etc) and how to enhance them.



Future work III

- Finally, it would be interesting to have a more detailed analysis of the costs associated to all this solution, from data gathering, processing and modeling to the production implementation itself. Also, taking into account other related factors, like infrastructure needed, legal aspects and ecological footprint [5] would give a more complete idea of the viability of global satellite image analysis.

Bibliography I

Thanks for your attention!

Peter Kareiva et al. "Domesticated Nature: Shaping Landscapes and Ecosystems for Human Welfare". In: *Science* 316.5833 (2007), pp. 1866–1869. ISSN: 0036-8075. DOI: [10.1126/science.1140170](https://doi.org/10.1126/science.1140170). eprint:
<https://science.sciencemag.org/content/316/5833/1866.full.pdf>. URL:
<https://science.sciencemag.org/content/316/5833/1866>.

Tim Newbold et al. "Global effects of land use on local terrestrial biodiversity". In: *Nature* 520 (Apr. 2015), 45 EP –. URL: <https://doi.org/10.1038/nature14324>.

Robert Costanza et al. "Changes in the global value of ecosystem services". In: *Global Environmental Change* 26 (2014), pp. 152 –158. ISSN: 0959-3780. DOI: <https://doi.org/10.1016/j.gloenvcha.2014.04.002>. URL:
<http://www.sciencedirect.com/science/article/pii/S0959378014000685>.

Kaiming He et al. "Deep Residual Learning for Image Recognition". In: *CoRR* abs/1512.03385 (2015). arXiv: 1512.03385. URL: <http://arxiv.org/abs/1512.03385>.

Emma Strubell, Ananya Ganesh, and Andrew McCallum. "Energy and Policy Considerations for Deep Learning in NLP". In: *arXiv e-prints*, arXiv:1906.02243 (2019), arXiv:1906.02243. arXiv: 1906.02243 [cs.CL].

Alert block

This is an alert bock

Alert block

This is another alert bock

