Degree Project Specification

Joar Gruneau
joar@gruneau.se

February 19, 2018

1 Formalities

Preliminary thesis title: Aerial image analysis with generative adversarial networks
Student's name: Joar Gruneau **Email:** joar@gruneau.se
Supervisor at CSC: Kevin Smith
Principal: BlackRock
Supervisor at principal's workspace:
Pascal Marcellini Pascal.Marcellini@blackrock.com

2 Background and objective

Convolutional neural networks have revolutionized the field of computer vision. The networks have a great generalization power and can be applied to many different types of tasks. There have been several papers that have used these network to detect and count objects in aerial images [5, 7, 12]. However detecting and counting objects in aerial images has proven to be a troublesome area. The objects of interest are usually very small compared to the image and there can be many objects in the same image. This causes the network to achieve bad performance if the entire image is fed in at once. To combat this some form of segmentation is required before the image is fed into the network. Earlier methods fed explicit image patches through the CNN [7]. This achieved good performance but at a great computational cost since redundant computation of low-level filters for overlapping patches had to be performed [8]. To combat this different forms of segmentation algorithms were used such as the mean-shift-algorithm which drastically decreased the number of patches which had to be fed through the network [3]. Today CNNs are used to do segmentation as well but it has proven difficult to construct a good loss function and semantic maps usually have to be post processed with conditional markov random fields (CRFs) and a morphological opening to ensure spatial connectivity [8] and object separation [5].

Goodfellow *et al* [6] proposed a generative adversarial network which is able to learn it's own loss function. It has later been shown that an adversarial loss combined with a cross entropy loss can lead to improved performance on several segmentation benchmarks [8]. Adversarial networks have also shown to have a great generalization power and have been especially successfully on small segmentation data sets [4, 11].

BlackRock has a team called AlphaGen which is responsible to construct models for making predictions of market trends. These trends are predicted by a model which uses many different signals as input. Fast analysis of aerial images could provide many signals of interest. If we for example over time are able to count the number of cars in a car park outside a supermarket, we can more accurately predict the sales of the same supermarket. This is only one application but there are several use cases where analysis of aerial images could give a better base to make predictions of economic trends.
The desired outcome of this this project is to investigate if a adversarial loss can improve the accuracy of the segmentor for segmentation of vehicles in aerial images.

3 Research question and method

This project aims to investigate if GANs be used to increase the performance of segmentation and object detection in aerial images. To answer this question the the proposed segmentation and object detection network will be evaluated both with and without a adversarial loss. Different types of adversarial losses will also be investigated such as patch-GAN and image-GAN since [11] has shown that the performance is closely coupled with the strength of the discriminator.

The generator will output several segmentation maps for the different classes. Depending on the accuracy of the segmentation maps the vehicles could either be detected by connected component extraction or a sliding regression network. The accuracy of these segmentation maps would first have to be determined before the proper next step could be chosen. The discriminator will then try to differentiate the generated bounding boxes from the ground truth ones. Different types of discriminators will be tested for the GAN. Image-GAN is the simplest which will determine at a image level if the generated boxes are true or false. Patch-GAN evaluates patches of the generated bounding boxes to determine if they are true or false. Since the discriminator usually wins the minimax game this means that we have some unused potential left in the discriminator. This project also desires to investigate if this potential can be harnessed by only training the generator on unlabelled

data until the discriminator no longer can differentiate between the generated and the ground truth images.

To test the performance of the methods the VEDAI dataset [10] will mainly be used since this is a well known benchmark dataset for aerial images. The set consists of 1271 colour aerial images with 1024*1024 resolution. The dataset has 10 different classes of vehicles together with their bounding boxes. The COWC vehicle dataset or the ISPRS semantic datasets of city landscape [1] could also be used.

4 Evaluation and news value

The evaluation of the objective will be compared to other methods on the VEDAI dataset as well as the own network without an adversarial loss. The metrics of the VEDAI dataset is the four common metrics, precision, recall, F1 score and average precision defined in [10]. The proposed networks will be evaluated on the first three metrics. As in earlier papers an intersection area divided by a union area (IoU) over 0.5 between the prediction and ground truth object box is considered a positive detection. The benefit of an adversarial loss for the segmentor will also be investigated. Here we will compare the pixel accuracy between the prediction and the ground truth.

At this writing moment there has not been a extensive study which has investigated image segmentation with GANs on aerial images. This project will therefore have a news value no matter of the outcome. Since GANs are such a new type of network this is still an area under development and most of the relevant papers came out in the end of last year. This project will add to this pool of growing knowledge and will therefore be of interest to people interested about GANs and image segmentation.

5 Pilot study

Since this type of network is so new there is no papers using GANs for aerial image segmentation and only a handful using them for for other types of segmentation. All of these papers came out in the end of last year. Therefore this pilot study needs gather knowledge about GANs applied to other image mapping tasks and convert that knowledge to this task. The GitHub repository [2] has a extensive collection of papers about GANs. There are a lot of papers on aerial image analysis about the VEDAI dataset which also will serve as an inspiration.

To decide if the second stage of the vehicle detection should use a connected component extraction or a sliding regression network the accuracy of the

probability maps have to be investigated. The pilot study will therefore be preformed in two parts where the first part will be about the GAN to generate the probability maps. After this fast evaluation will be performed and the second part about connected component extraction or sliding window regression depending on the outcome of the this.

6 Conditions and schedule

As mentioned above the VEDAI dataset will be used which is open source. The networks will be trained on a Amazon EC2 spot instance. I will have continuous contact with the external supervisor during the project. He has promised to also serve as a interface towards the BlackRock AlphaGen team if I had some more technical problems he was not be able to help with.

6.1 Schedule

This is the preliminary layout of the schedule. Remember that the thesis writing will be done in parallel so the duration column does not specify full time work.

Task	Start date	End date	Duration
Preparations	5/2	9/2	4 d
Pilot study part 1	24/1	16/2	3.5 w
Investigating GAN models	24/1	14/2	3 w
Investigating transfer learning for GANs	14/2	16/2	3 d
Short evaluation	19/2	28/2	1.5 w
Process dataset	19/2	21/2	3 d
Constructing and training simple prototype GAN	21/2	28/2	1 w
Pilot study part 2	28/2	7/3	1 w
Implementation	7/3	11/4	5 w
Evaluation of results	11/4	18/4	1 w
Thesis writing	24/1	6/5	13.5 w
Background and related work	24/1	7/3	6 w
Methodology	7/3	11/4	5 w
Results and conclusion	11/4	6/5	3.5 w
Additional research	6/5	20/5	2 w
Final preparations	20/5	4/6	2 w

6.2 Schedule definitions

The major parts of the schedule are:

1. **Preparations** Gaining broad knowledge about the project area and writing the project description.

2. **Pilot study part 1** Study of GANs to produce the probability map.

3. **Short evaluation** Short evaluation to determine the next step of the project.

4. **Pilot study part 2** Study of the selected method to perform vehicle detection.

5. **Implementation** Implementing the full solution.

6. **Evaluation of results** This phase concerns evaluating the proposed method on the given datasets.

7. **Thesis writing and supplementary studies** Writing the thesis is done in parallel with the pilot study and implementation. The background and related work section will be written in parallel to the pilot study and the methodology section will be written in parallel with the implementation.

8. **Additional research** Time for additional research if needed or otherwise spent perfecting the report.

9. **Final preparations** Preparing for presentation and opposition.

References

[1] 2d Semantic Labeling - ISPRS, . URL `http://www2.isprs.org/commissions/comm3/wg4/semantic-labeling.html`.

[2] GitHub - nightrome/really-awesome-gan: A list of papers on Generative Adversarial (Neural) Networks, . URL `https://github.com/nightrome/really-awesome-gan`.

[3] N. Ammour, H. Alhichri, Y. Bazi, B. Benjdira, N. Alajlan, and M. Zuair. Deep Learning Approach for Car Detection in UAV Imagery. *Remote Sensing*, 9(4):312, Mar. 2017. doi: 10.3390/rs9040312. URL `http://www.mdpi.com/2072-4292/9/4/312`.

[4] A. Arbelle and T. R. Raviv. Microscopy Cell Segmentation via Adversarial Neural Networks. *arXiv:1709.05860 [cs]*, Sept. 2017. URL `http://arxiv.org/abs/1709.05860`. arXiv: 1709.05860.

[5] N. Audebert, B. Le Saux, and S. Lefèvre. Segment-before-Detect: Vehicle Detection and Classification through Semantic Segmentation of Aerial Images. *Remote Sensing*, 9(4):368, Apr. 2017. doi: 10.3390/rs9040368. URL `http://www.mdpi.com/2072-4292/9/4/368`.

[6] I. Goodfellow. NIPS 2016 Tutorial: Generative Adversarial Networks. *arXiv:1701.00160 [cs]*, Dec. 2016. URL `http://arxiv.org/abs/1701.00160`. arXiv: 1701.00160.

[7] A. C. Holt, E. Y. W. Seto, T. Rivard, and P. Gong. Object-based detection and classification of Vehicles from high-resolution aerial photography. *Photogrammetric Engineering and Remote Sensing*, 75(7):871–880, July 2009. ISSN 0099-1112. URL `https://iths.pure.elsevier.com/en/publications/object-based-detection-and-classification-of-vehicles-from-high-r`.

[8] P. Luc, C. Couprie, S. Chintala, and J. Verbeek. Semantic Segmentation using Adversarial Networks. *arXiv:1611.08408 [cs]*, Nov. 2016. URL `http://arxiv.org/abs/1611.08408`. arXiv: 1611.08408.

[9] T. N. Mundhenk. Cars Overhead With Context Dataset at LLNL. URL `http://gdo-datasci.ucllnl.org/cowc/`.

[10] S. Razakarivony and F. Jurie. Vehicle Detection in Aerial Imagery : A small target detection benchmark. *Journal of Visual Communication and Image Representation, Elsevier*, Mar. 2015. URL `https://hal.archives-ouvertes.fr/hal-01122605`.

[11] J. Son, S. J. Park, and K.-H. Jung. Retinal Vessel Segmentation in Fundoscopic Images with Generative Adversarial Networks.

arXiv:1706.09318 [cs], June 2017. URL `http://arxiv.org/abs/1706.09318`. arXiv: 1706.09318.

[12] J. Zhong, T. Lei, and G. Yao. Robust Vehicle Detection in Aerial Images Based on Cascaded Convolutional Neural Networks. *Sensors (Basel, Switzerland)*, 17(12), Nov. 2017. ISSN 1424-8220. doi: 10.3390/s17122720. URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5751529/`.