University of Münster
Institute for Geoinformatics

Proposal

„Automatic detection of vehicles in multispectral aerial images using deep learning“

First reviewer: Prof. Dr. Benjamin Risse
Second reviewer: Dr. Christian Knoth
Release date: tbd.
Deadline: tbd.

Submitted by: Timo Lietmeyer
Born on: 23.05.1999
E-mail address: timolietmeyer@uni-muenster.de
Student ID number: 459 169
Study program: M. Sc. Geoinformatics and Spatial Data Science

Motivation und Kontext

Motivation and Context

The easy availability of high-resolution satellite and aerial images makes it possible to detect objects worldwide. Furthermore, the high computing power required for machine learning is becoming increasingly cheaper, making the application of artificial intelligence for fully automated analysis easier. The monitoring of destroyed (military) equipment or building structures in conflict regions, such as Darfur [1] or Ukraine, can be supported by artificial intelligence. Aerial imagery can be used well here because the security situation on the ground in conflict regions can make it dangerous to take photos on the ground.
Monitoring military equipment at national borders could be used to generate a more accurate assessment of the situation to determine whether a conflict is imminent. It is also possible to support the work of human analysts, as the models provide suggestions for possible detections and classifications of objects.
Machine learning can be applied to automate the process of processing larger data sets. One challenge here is training the algorithms, since the availability of training data is limited.

One application scenario would be to evaluate human mobility behavior when counting cars in large public parking lots. Here, traffic volumes can also be tracked nationwide by the AI if vehicles can be detected on a satellite image. This approach can be used to estimate whether more parking spaces are needed or whether the number of current spaces is sufficient. Further questions would be whether it is possible to distinguish between moving and stationary cars.

Furthermore, this work can evaluate whether the precision of object detection and classification of vehicles in electro-optical, multi-spectral aerial images can be improved by using more image channels (NIR, IR). This can be done by comparing the results when training the same deep learning model with only 3 and once with more channels. Since vehicles appear relatively small on high-resolution satellite images, it is also possible to evaluate how well a deep learning model is suited to classifying very small objects, or whether an additional channel improves precision.

Research Questions

The following research questions can be answered in the context of this work:

- Can object detection and classification on multi-spectral aerial images be improved by using four or more channels?
- Is deep learning suitable for detecting very small objects on high-resolution aerial and/or satellite images?
- How does the IR / NIR band help to classify and detect objects on high-resolution satellite images?

Literature review

Aerial or satellite images can already be used to detect mass graves [2]. This approach can be applied in near real-time or in the past [2]. It is based on satellite images that consist of different bands and can also be used to detect secret mass graves [2]. Here it could be possible to modify it to detect mass graves in Ukraine [2].
Other approaches include the detection of individual graves using airborne imagery [3].
In addition, satellite images of Chinese crematoria were used to verify the official government data on Covid mortality. This, in combination with eyewitness interviews, was used to cast doubt on the official mortality figures [4].

Satellite images have already been used to detect the destruction of hut structures in Darfur [1]. Methods have been developed to detect and count these structures [5]. This may also be possible to apply to vehicles [5].
High-resolution satellite images can also be used with artificial intelligence to analyze poverty [6].

RGB images from drones are sufficient for object detection in most scenarios. Thermal (IR) images can extend the possibilities for object detection at night or for hidden objects. One problem is the lack of available training data for IR images from drones. There have been approaches to mitigate this problem by creating synthetic IR images [7].

Methods

The 'You Only Look Once' (YOLO) algorithm is ideal for detecting objects on high-resolution satellite images. This algorithm, which was released in version 9 in 2024, enables fast and accurate object detection. YOLO is based on a single neural network architecture that divides the image into grids and makes predictions about the position and class of objects for each grid.

The main advantages of YOLO are its speed and accuracy, which make it possible to efficiently analyze large data sets. For this work, YOLO is used to detect and classify vehicles on satellite images. Additional multispectral channels such as NIR or IR could be integrated to allow a comparison to the pure 3-channel RGB training. Comparability can be ensured if the same training and test data is used, where the only difference is the additional band.

Data basis

The 'Vehicle Detection in Aerial Images' (VEDAI) dataset [8] from 2015 is suitable as a data basis because it contains high-resolution aerial images that are specifically suitable for vehicle detection [9]. It includes annotated data showing vehicles in different scenarios, sizes and orientations. It is also a benchmark for the detection of very small objects. The dataset contains both RGB images and multispectral data, making it ideal for investigating the effects of additional channels such as NIR or IR on object detection. There are more than 3700 objects annotated in over 1200 images. These objects are grouped into 8 classes (Boat, Camping Car, Car, Pickup Plane, Tractor, Truck, Van and

Others) and the background of the objects is varied, which increases the robustness of the trained model.
The resolution of the images is 12.5 cm $\times$12.5 cm per pixel, which is sufficient to recognize individual vehicles. It may be useful to downscale to 30 cm per pixel to imitate the resolution of satellite images.

Data preparation includes steps such as dividing the images into training and test data, as well as preparing the annotations for the YOLO algorithm. After these steps, the training can take place on the high-performance cluster (HPC) PALMA at the University of Münster. Based on the result matrix, a comparison of the two models will then be made, since the only difference is the additional band.

Expected results

The aim of the work is to evaluate whether adding another image channel (such as NIR or IR) improves the precision and prediction quality of the deep learning model YOLOv9. It is possible that adding the image channel will decrease or increase the accuracy because adding information affects the prediction quality.
In summary, the goal of this work can be an assessment of whether deep learning on multi-channel aerial or satellite images can be improved by adding more information.

Bibliography

[1] Christian Knoth and Edzer Pebesma. Detecting dwelling destruction in darfur through object-based change analysis of very high-resolution imagery. *International Journal of Remote Sensing*, 38:273–295, 1 2017. ISSN 13665901. doi: 10.1080/01431161.2016.1266105. URL `https://www.tandfonline.com/doi/abs/10.1080/01431161.2016.1266105`.

[2] M. Kalacska and L.S. Bell. Remote sensing as a tool for the detection of clandestine mass graves. *Canadian Society of Forensic Science Journal*, 39:1–13, 1 2006. ISSN 0008-5030. doi: 10.1080/00085030.2006.10757132.

[3] G. Leblanc, M. Kalacska, and R. Soffer. Detection of single graves by airborne hyperspectral imaging. *Forensic Science International*, 245:17–23, 12 2014. ISSN 0379-0738. doi: 10.1016/J.FORSCIINT.2014.08.020.

[4] Corona in china: Satellitenbilder von krematorien deuten auf deutlich mehr tote hin - der spiegel, 2023. URL `https://www.spiegel.de/ausland/corona-satellitenbilder--chinesischer-krematorien-deuten-auf-deutlich-mehr-tote-hin-a-c4305852-d092-4e54--a210-c88b820d564c`.

[5] Thomas Kemper, Malgorzata Jenerowicz, Martino Pesaresi, and Pierre Soille. Enumeration of dwellings in darfur camps from geoeye-1 satellite images using mathematical morphology. *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, 4, 2011. doi: 10.1109/JSTARS.2010.2053700. URL `http://ieeexplore.ieee.org`.

[6] Ola Hall, Francis Dompae, Ibrahim Wahab, and Fred Mawunyo Dzanku. A review of machine learning and satellite imagery for poverty prediction: Implications for development research and applications. *Journal of International Development*, 35:1753–1768, 10 2023. ISSN 1099-1328. doi: 10.1002/JID.3751. URL `https://onlinelibrary.wiley.com/doi/full/10.1002/jid.3751` `https://onlinelibrary.wiley.com/doi/10.1002/jid.3751`.

[7] Lizhi Yang, Ruhang Ma, and Avideh Zakhor. Drone object detection using rgb/ir fusion, 2022. URL `https://arxiv.org/abs/2201.03786`.

[8] Vehicle detection in aerial imagery (vedai) : a benchmark. URL `https://downloads.greyc.fr/vedai/`.

[9] Sébastien Razakarivony, Frédéric Jurie, Sebastien Razakarivony, and Frederic Jurie. Vehicle detection in aerial imagery : A small target detection benchmark. *Journal of Visual Communication and Image Representation*, 2015. URL `https://hal.science/hal-01122605v2`.