MapAI Competition Submission for Team Kaborg

Karl Audun Borgersen¹, Morten Grundetjern² E-mail any correspondence to:
karl.audun.borgersen@uia.no

1. Centre for Artificial Intelligence Research, Department of ICT, University of Agder, Grimstad, Norway
2. Centre for Artificial Intelligence Research, Department of ICT, University of Agder, Grimstad, Norway

Abstract
This paper constitutes the final portion of our submission to the MapAI image segmentation competition [1]. We introduce a Stable Diffusion (SD) [2] model with a particular emphasis on augmenting image segmentation datasets. Due to an unfortunate mistake made during the training of our model iterations, the benefits of these techniques are not reflected in the final submission score. We reach only a combined score of 0.62. Nevertheless, the method shows potential and can be applied to accurate image segmentation in future projects.

Keywords: artificial intelligence; map segmentation; deep learning; dataset augmentation

Introduction
The task of segmenting building from map data provides a series of interesting challenges. This paper discusses one possible approach to tackling these challenges within the context of a contest held by Sander Jyhne in collaboration with the Norwegian Artificial Intelligence Research Consortium (NORA) along with the Centre for Artificial Intelligence Research at the University of Agder (CAIR), the Norwegian Mapping Authority, AI:Hub, Norkart, and The Danish Agency for Data Supply and Infrastructure. The competition is broken into two separate tasks. Task 1 involves the segmentation of buildings within aerial real-world photographs. The second task also involves building segmentation but is based on using LIDAR data.

Materials and methods
An iteration of U-net [3] was used as the primary segmentation model for task 1 and task 2 of this competition. This U-Net model used Res-Net 50 [4] as its backbone model. We observed significant increases in scores when increasing the batch size and decreasing the learning rate on each task to around $1 \times 10^{-5}$. The loss used for the U-Net model was Dice Loss, which further improved the score of the model.

Stable-Diffusion-Based Dataset Augmentation
An approach commonly used for improving the score of most machine learning models is to introduce a degree of augmentation to the original dataset to inflate the total amount of accessible training data. Such augmentations are traditionally just simple rules-based changes to the original image format, such as rotation, cropping, the introduction of noise, or color channel alterations. The release of diffusion models to the public, however, might permanently introduce a new alternative. The generation of entirely new image features. In the case of, for instance, more traditional image classification tasks, such as ImageNet [5], this can take the form of generating images based on the labels provided. However, the results of such augmentation are unlikely to outperform the classification method used for the image generation model itself. But in the case of segmentation challenges, we can use image generation to significantly alter the context of the label data while leaving the actual label data unchanged, or vice versa. Figure 1 displays the result of augmenting the ground around a label, and Figure 2 displays the results of repainting the original roofs instead.

As shown in Figure 3, the SD 1.5 inpainting model is better at replacing missing context in a vacuum. However, it is not capable of working off of the features originally present in the paper. It will often, therefore, introduce features that should be labeled as rooftops but are not.

This approach to augmentation was only used for Task 1 of the competition.

© 2023 Author(s). This is an open access article licensed under the Creative Commons Attribution License 4.0. (http://creativecommons.org/licenses/by/4.0/)
Results

Table 1, 2, and 3 present the evaluation results for task 1, task 2, and the final evaluation score, respectively. Our solution scored 9th place in this competition.

Table 1: Evaluation results for task 1

<table>
<thead>
<tr>
<th>IoU T1</th>
<th>BIoU T1</th>
<th>Total T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7112</td>
<td>0.5195</td>
<td>0.6154</td>
</tr>
</tbody>
</table>

Table 2: Evaluation results for task 2

<table>
<thead>
<tr>
<th>IoU T2</th>
<th>BIoU T2</th>
<th>Total T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6890</td>
<td>0.5616</td>
<td>0.6253</td>
</tr>
</tbody>
</table>

Table 3: Final score for the evaluation results

<table>
<thead>
<tr>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6203</td>
</tr>
</tbody>
</table>

Discussion

We did see a slight increase in score from using SD augmentation on the scale of <0.01 with some trial and error. The best results were achieved by augmenting both the roofs and the ground. Then, as the model got closer and closer to its best score, the augmented data was slowly added as a larger and larger part of the training dataset. The score increase achieved with this technique was less than originally hoped. A significant portion of the calculated score on both datasets is made up of the BIoU metric, which measures the degree to which the model is capable of tracking the border of the labels. This is, predictably, where machine learning models struggle the most. Future work using this method of augmentation might do well by focusing on augmenting the border areas around each label rather than the entire image.

The actual final scores of this competition would likely have benefitted far more from focused hyperparameter tuning and better model choices than the methods described within this submission paper.

The final evaluation score was hampered by the accidental presence of an augmentation technique that altered the format of both the training and validation datasets but not the evaluation dataset of the competition. The presence of this mistake is likely to have reduced the effectiveness of the methods described in this paper as well.

Conflict of interest

Authors state no conflict of interest.
References


