# Boosting Instance Segmentation with Synthetic Data - Supplementary Material

Florentin Poucin
École des Ponts Paristech
Champs-sur-Marne, France
florentin.poucin@eleves.enpc.fr

Andrea Kraus, Martin Simon
Valeo Schalter und Sensoren GmbH,
Kronach, Germany
{andrea.kraus, martin.simon}@valeo.com

While the aim of the main paper is to improve the global performance of the network using synthetic data, in this additional content, we explore the ability of synthetic images to improve the network precision for a specific instance category. We focus on the less represented class among Cityscapes images: "trains".

## Experiments with class targeting

The idea is to use features related to "trains" from synthetic data, in order to improve the global network comprehension and precision on this specific class, on both real and synthetic images. Figure 1 is composed of Synscapes images containing different instances of "trains" in very diverse situations. These instances are more or less occluded, more or less deep in the image and the instance density around them varies a lot. However, the instances themselves are quite similar. The shapes and colors are globally the same and some elements like the train pantograph are very specific in these synthetic images. On the one hand, the Synscpaes data set shows a wide variety of situations, but on the other hand, it shows a lack of diversity regarding train characteristics.

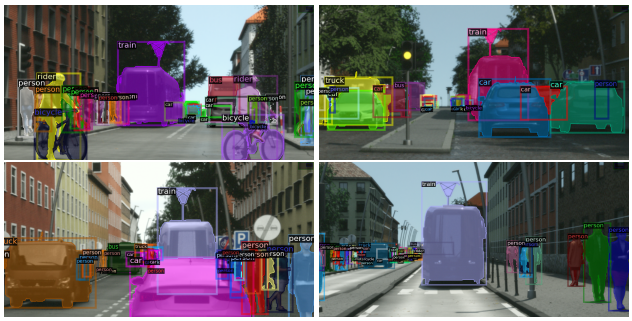

Figure 1. **Train instances from the Synscapes data set.**

We explore class targeting by evaluating the impact of training image properties, regarding instances of "train", on the final network performance. We create four different mixed data sets. Each of them is composed of 8,975 images, including the 2,975 Cityscapes training images to which we add 6,000 synthetic images with different properties related to instances of "train":

- Synthetic set 1 : Synscapes images containing no instance of "trains";
- Synthetic set 2 : Synscapes images randomly selected without special property;
- Synthetic set 3 : Synscapes images with an over representation of "trains" label, selected according to the count of instances corresponding to "trains";
- Synthetic set 4 : Synscapes images with an over representation of "trains" label, selected according to the count of pixels corresponding to "trains";

| Training data set | "trains" AP | Mean AP |
|---|---|---|
| Cityscapes | 36,3 | 36,2 |
| Cityscapes + Synthetic set 1 | 41,3 | 38,3 |
| Cityscapes + Synthetic set 2 | 43,0 | 38,9 |
| Cityscapes + Synthetic set 3 | **47,3** | **39,3** |
| Cityscapes + Synthetic set 4 | 46,8 | 39,1 |

Table 1. **Impact of training data set properties on Average Precision scores.**

The network is trained of the different data sets for 50 epochs and its performance are summarized in Table 1. First of all, we can observe an improvement of the precision on "trains" for each mixed data set compared to Cityscapes only. Even though the synthetic images do not include instance of "trains", they make the network precision better. Nevertheless, synthetic features of "trains" additionally improves the network. The best precision is obtained when "trains" are over represented in terms of number of instances. This label targeting increases the precision on "trains" by more than 10% and also increases the global average precision, even compared with mixed data sets with the same number of images. Class targeting is a good option to improve the network for a specific task, but also for an overall improvement of performance.