Impact of Data Anonymization on Semantic Segmentation

Jingxing Zhou

KI DELTA LEARNING
Scalable AI for Automated Driving

Motivation

- Strict data privacy regulations, e.g., GDPR, requires anonymization
- Datasets are essential for the development of data driven models
- Impact of using anonymized training data within automotive domain (ADAS/HAD) remains unclear

Experimental Setup

- Train the neural networks with Cityscapes dataset [1] utilizing 4 anonymization patterns: Crop Out, Gaussian Filter, Random Crop, Image Resynthesizing [2].
- Validate and test on original un-anonymized dataset, which is analogous to real driving scenarios.
- Statistical analysis using the Wilcoxon signed-rank and the Mann-Whitney U test.

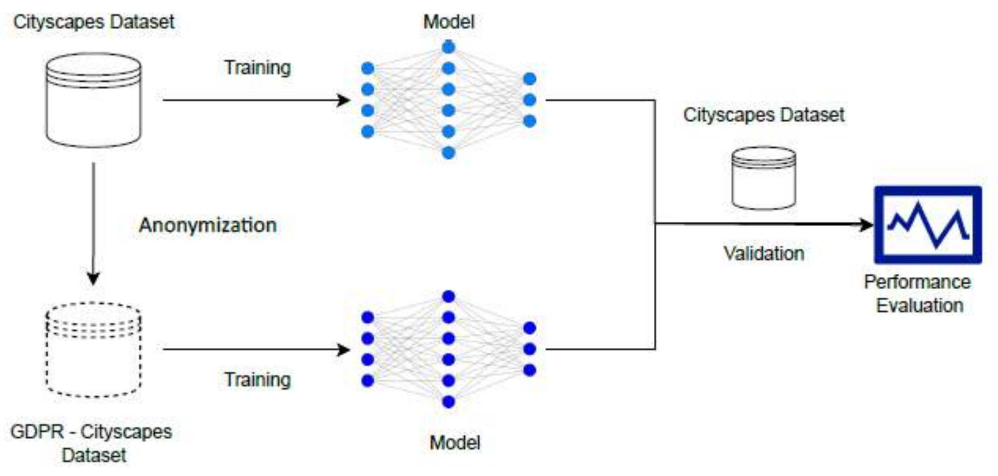

Figure 1: Experimental Setup. In order to simulate real driving situations, where the image stream is directly fed to the computing unit, we run the validation by applying the trained model with on the original Cityscapes validation dataset.

Quantitative Results

	Crop Out	Random	Gaussian	Resynthesis
Class				
Person	**-0.99%**	-0.93%	-0.43%	0.08%
Rider	**-2.87%**	-2.00%	-0.82%	0.09%
Car	**-0.29%**	-0.14%	-0.17%	-0.13%
Bicycle	-0.22%	**-0.56%**	-0.40%	-0.09%
Category				
Human	**-0.89%**	-0.83%	-0.40%	0.07%
Vehicle	-0.07%	**-0.18%**	-0.10%	-0.01%
mIoU	**-0.87%**	0.02%	-0.31%	-0.05%

Table 1: Relative IoU Changes w.r.t. Anonymization Pattern

Table 1 shows the average IoU change caused by deploying the given anonymization pattern in the training dataset. The Crop Out method has the highest IoU loss. The application of the Gaussian filter results in a moderate change. In contrast, image resynthesis does not have a noticeable negative impact on the image segmentation. Besides that, we observe a decrease in accuracy for classes that are not directly relevant to the anonymization, like *Bicycle*. Category-wise, image anonymization has a much more noticeable impact on the classes *Person* and *Rider* in category *Human* compared to *Car* in category of *Vehicle*.

Additionally, we found out that models trained with high resolution input images suffer more from image data anonymization in all classes than trained with low resolution images. Networks that learns multi-scale features, can better cope with anonymized images. The performance limitations caused by image data anonymization decrease with the increasing backbone capacity.

Quantitative Results

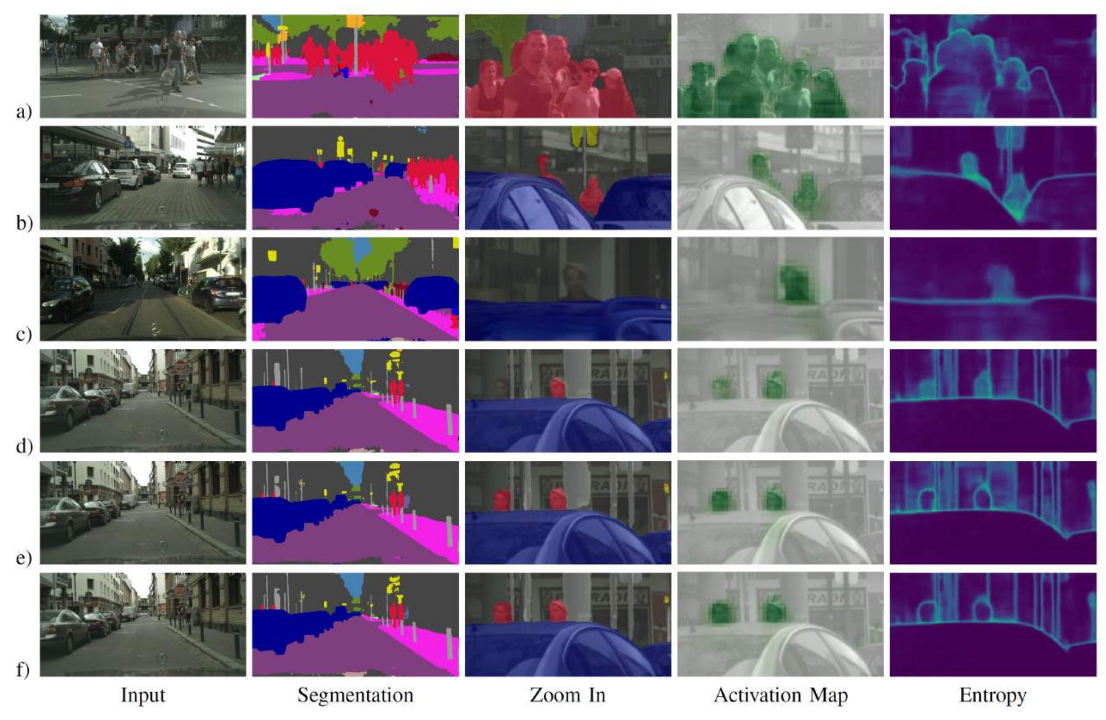

Figure 2: Qualitative results on Cityscapes test dataset from FCN network with ResNet 18 backbone trained on gaussian anonymized Cityscapes dataset.

Figure 2 demonstrates qualitative results generated from the model, where Row a) shows the scenario where the model is able to cope with dense crowds, when the pedestrians are not obstructed from other objects. Row b) to d) are some scenarios that the network trained by anonymized dataset failed to detect pedestrians and cyclist, while Row e) and f) present the segmentation results with low uncertainty from the same model trained with original and re-synthesised dataset.

Conclusion

- Better generalization and image segmentation with complex backbones and multi-scale features.
- Face anonymization degrades performance on obstructed *pedestrians* and *riders.*
- Marginal effect on the class *Vehicle.*
- Generative models help to mitigate the side effect of image data anonymization.

Reference

[1] Cordts, Marius, et al. "The cityscapes dataset for semantic urban scene understanding." Proceedings of CVPR. 2016.
[2] Hukkelås, Håkon, et al. "DeepPrivacy: A generative adversarial network for face anonymization." ISVC. 2019.

Real World Robustness

Partners

External partners

For more information contact:
jingxing.zhou@porsche-engineering.de

KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

KI FAMILIE

www.ki-deltalearning.de @KI_Familie KI Familie

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag