Instance-wise Uncertainty for Class Imbalance in Semantic Segmentation - Supplementary Material

The present document reports all training preprocessing steps and chosen hyperparameters in detail. Code is available at `https://github.com/luislcc/instance_uncertainty`.

1 Preprocessing

All input images and ground truth segmentation masks were resized to 768x768 pixels using bilinear and nearest neighbor interpolation, respectively. No data augmentation techniques were employed. Image normalization was performed using the ImageNet means and standard deviations.

2 Training

All DeepLabV3+ models are composed by a backbone followed by a classifier. The learning rate of the backbone was always set to be 0.1 times the classifier learning rate.

2.1 DeepLabv3+ with ResNet50 and ResNet101 Training Hyperparameters

- Learning Rate: 1e-4
- Output Stride: 16
- Batch Size: 2
- Validation Batch Size: 1
- Epochs: 60

2.2 DeepLabv3+ with MobileNetV2 Hyperparameters

- Learning Rate: 1e-3
- Output Stride: 16

- Batch Size: 2
- Validation Batch Size: 1
- Epochs: 60

The Adam optimizer was used for all experiments with Pytorch's default parameters. A polynomial learning rate scheduler was used with the power parameter set to 0.9. Each model is evaluated on the validation set every 100 iterations and the best performing model is saved.

These parameters were inaltered and were the same in all experiments.

3 Full Tables for Experiment 1

Table 1: Performance of DeepLabV3+ on all classes of the ACDC Dataset. These values were obtained on the respective validation set.

	Standard Training (**Baseline**)			Instance Uncertainty Training		
Class	MobileNet	ResNet50	ResNet101	MobileNet	ResNet50	ResNet101
Road	0.94±0.00	0.95±0.00	0.95±0.00	0.94±0.00	0.94±0.00	0.94±0.00
Sidewalk	0.77±0.01	0.79±0.01	0.79±0.00	0.76±0.01	0.76±0.00	0.76±0.00
Building	0.82±0.00	0.83±0.00	0.84±0.00	0.82±0.00	0.83±0.00	0.83±0.00
Wall	0.50±0.01	0.51±0.01	0.52±0.01	0.49±0.01	0.50±0.01	0.51±0.01
Fence	0.40±0.01	0.42±0.01	0.42±0.01	0.39±0.01	0.40±0.01	0.42±0.01
Pole	0.49±0.01	0.51±0.01	0.51±0.01	0.49±0.01	0.50±0.01	0.51±0.01
Traffic Light	0.59±0.01	0.63±0.01	0.64±0.01	0.59±0.01	0.64±0.01	0.65±0.01
Traffic Sign	0.47±0.01	0.53±0.01	0.53±0.01	0.48±0.01	0.51±0.01	0.53±0.01
Vegetation	0.84±0.00	0.84±0.00	0.85±0.00	0.84±0.00	0.84±0.00	0.85±0.00
Terrain	0.47±0.01	0.49±0.01	0.49±0.01	0.47±0.01	0.46±0.01	0.48±0.01
Sky	0.95±0.00	0.95±0.00	0.95±0.00	0.94±0.00	0.95±0.00	0.95±0.00
Person	0.40±0.02	0.42±0.02	0.43±0.02	0.42±0.02	0.45±0.01	0.49±0.01
Rider	0.20±0.07	0.15±0.06	0.16±0.05	0.22±0.06	0.20±0.02	0.23±0.05
Car	0.80±0.06	0.82±0.01	0.83±0.01	0.80±0.01	0.82±0.01	0.82±0.01
Truck	0.37±0.03	0.41±0.05	0.49±0.04	0.39±0.03	0.46±0.03	0.48±0.02
Bus	0.74±0.02	0.74±0.03	0.77±0.03	0.73±0.03	0.71±0.04	0.69±0.02
Train	0.82±0.01	0.85±0.01	0.85±0.01	0.81±0.01	0.84±0.01	0.83±0.01
Motorcycle	0.24±0.04	0.25±0.03	0.30±0.04	0.26±0.03	0.29±0.03	0.36±0.02
Bicycle	0.38±0.03	0.36±0.04	0.35±0.04	0.40±0.04	0.37±0.03	0.41±0.04
mIoU	0.59	0.60	0.61	0.59	0.60	0.62

Table 2: The class IoU performance values of both Standard and Instance Uncertainty training on the CityScapes validation set. These are IoU values for DeepLabV3+ models with a ResNet101 backbone

Class	Standard Training (**Baseline**)	Instance Uncertainty Training
Road	0.98 ±0.00	0.98 ±0.00
Sidewalk	0.81 ±0.00	0.81 ±0.00
Building	0.89 ±0.00	0.89 ±0.00
Wall	0.52 ±0.02	0.52 ±0.02
Fence	0.48 ±0.01	0.48 ±0.01
Pole	0.43 ±0.01	0.44 ±0.01
Traffic Light	0.51 ±0.01	0.52 ±0.01
Traffic Sign	0.64 ±0.01	0.65 ±0.01
Vegetation	0.90 ±0.00	0.90 ±0.00
Terrain	0.61 ±0.01	0.61 ±0.01
Sky	0.93 ±0.00	0.93 ±0.00
Person	0.69 ±0.00	0.70 ±0.00
Rider	0.46 ±0.01	0.47 ±0.01
Car	0.92 ±0.00	0.92 ±0.00
Truck	0.70 ±0.02	0.68 ±0.02
Bus	0.77 ±0.01	0.77 ±0.01
Train	0.66 ±0.03	0.65 ±0.03
Motorcycle	0.53 ±0.01	0.53 ±0.01
Bicycle	0.66 ±0.00	0.67 ±0.00
mIoU	0.69	0.69

4 Full Table for Experiment 2

Table 3: Performance of DeepLabV3+ with ResNet101 backbone on all classes. These values were obtained on the respective validation set. CS → ACDC denotes "trained on CityScapes (CS) and tested on ACDC validation set".

	Standard Training (**Baseline**)		Instance Uncertainty Training	
Datasets	CS → ACDC	ACDC → CS	CS → ACDC	ACDC → CS
Road	0.79±0.01	0.85±0.05	0.79±0.01	0.90±0.02
Sidewalk	0.35±0.04	0.47±0.04	0.34 ±0.03	0.56±0.02
Building	0.60±0.03	0.76±0.03	0.60±0.03	0.82±0.01
Wall	0.24±0.02	0.11±0.03	0.24±0.02	0.21±0.02
Fence	0.23±0.02	0.17±0.02	0.23±0.02	0.23±0.01
Pole	0.32±0.02	0.27±0.01	0.34±0.02	0.31±0.01
Traffic Light	0.44±0.02	0.21±0.02	0.46±0.02	0.31±0.01
Traffic Sign	0.44±0.03	0.41±0.01	0.30 ±0.04	0.47±0.01
Vegetation	0.72±0.01	0.81±0.02	0.72±0.01	0.83±0.01
Terrain	0.26±0.01	0.15±0.07	0.27±0.01	0.26±0.04
Sky	0.76±0.01	0.78±0.03	0.75±0.04	0.82±0.01
Person	0.23±0.04	0.25±0.04	0.24±0.04	0.52±0.01
Rider	0.14±0.05	0.07±0.02	0.12±0.04	0.20±0.02
Car	0.67±0.05	0.80±0.03	0.66±0.05	0.86±0.00
Truck	0.30±0.05	0.18±0.03	0.29±0.05	0.40±0.03
Bus	0.25±0.04	0.22±0.05	0.22±0.05	0.38±0.03
Motorcycle	0.40±0.04	0.07±0.03	0.39±0.03	0.17±0.06
Train	0.24 ±0.03	0.06±0.02	0.24±0.03	0.15 ±0.03
Bicycle	0.19 ±0.03	0.37±0.01	0.39±0.01	0.54±0.01
mIoU	0.39±0.01	0.37±0.01	0.39±0.01	0.47±0.01