Supplemental Material

We provide further information about the utilized image corruptions and the conducted experiments. In more detail, we first show examples of every image corruption, and we give further details of our proposed image corruptions (section A). To make the image corruptions mutually comparable, we provide the Signal-to-Noise ratio for image corruptions of category noise (section A.3).
In section B, we provide supplementary information about the experimental setup (section B.1, B.2), we explain the difference of the utilized evaluation metrics (*i.e.*, CD and rCD) in more detail (section B.3), we discuss possible causes of the effect of architectural design choices (section B.4), and we show qualitative results (section B.5).
In addition, we report the individual evaluation metric scores (*i.e.*, mIoU, CD, and rCD) for Cityscapes (section B.6), PASCAL VOC 2012 (section B.7), and ADE20K (section B.8). We further show the performance of many models for different severity levels of many image corruptions (section B.10).

A. Image Corruption Models

A.1. ImageNet-C

In this section, we illustrate the used image corruptions of ImageNet-C[1]. Figure A.1 shows the image corruptions of the categories blur, noise, digital, and weather of ImageNet-C. To make the image corruption clearly visible, we selected each example of severity level three or higher. The Figure is best viewed in color.

A.2. Proposed Image Corruptions

In this section, we provide more details about our image corruptions, *i.e.*, the proposed image noise model, PSF blur, and geometric distortion. Figure A.3 shows examples of our proposed image corruptions.

Intensity-Dependent Noise Model. In the main paper, we proposed a noise model that incorporates intensity-dependent chrominance und luminance noise components that are both added to original pixel intensities in linear color space. Here, the term *chrominance noise* means that a random noise component for an image pixel is drawn for each color channel independently, resulting thus in color noise. *Luminance noise*, on the other hand, refers to a random noise value that is added to each channel of a pixel equally, resulting hence in gray-scale noise. We model the noisy pixel intensity for a color channel c as a random variable $I_{noise,c}$:

$$\begin{aligned} I_{noise,c}(\Phi_c, N_{luminance}, N_{chrominance,c}; w_s) = \\ log_2(2^{\Phi_c} + w_s \cdot (N_{luminance} + N_{chrominance,c})) \end{aligned} \tag{1}$$

where Φ_c is the normalized pixel intensity of color channel c, $N_{luminance}$ and $N_{chrominance}$ are random variables following a Normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$, w_s is a weight factor, parameterized by severity level s.

PSF blur. Every optical system, *e.g.*, the lens array of a camera, exhibits optical aberrations. Many of them cause image blur. Point-spread-functions aggregate every optical aberration that results in blur. The point-spread-functions of an optical system are typically spatially-varying, meaning, for example, that the degree of blur is at the image edge more pronounced than in the center of the image. Figure A.4 illustrates the intensity distribution of a PSF kernel, where most of its energy is punctually centered.

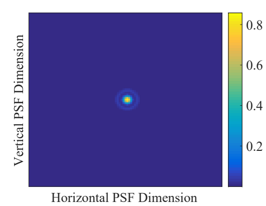


Figure A.4: The normalized intensity distribution of a PSF kernel of our proposed PSF blur.

Figure A.2 illustrates the intensity distribution of several PSF kernels utilized in the main paper. Each row corresponds to a specific PSF blur kernel at the respective angle of incidence, *i.e.*, the higher the angle of incidence, the higher the distance to the image center. Note that the PSF kernel varies its shape within a severity level (*i.e.*, column). The intensity of a PSF kernel is spatially more distributed for higher severity levels.

Geometric distortion. Distortion parameters of an optical system vary over time, are affected by environmental influences, differ from calibration stages, and thus, may never be fully compensated. Additionally, image warping may introduce re-sampling artifacts, degrading the informational content of an image. It can hence be preferable to utilize the original (i.e., geometrically distorted) image [3, p.192f]. We used the command-line tool *ImageMagick* to apply a radially-symmetric barrel distortion as a polynomial of grade 4 to both the RGB and ground-truth images. It is essential to use the nearest-neighbor filter for color determination of the ground truth, as otherwise, the class labels are corrupted.

[1]https://github.com/hendrycks/robustness/tree/master/ImageNet-C

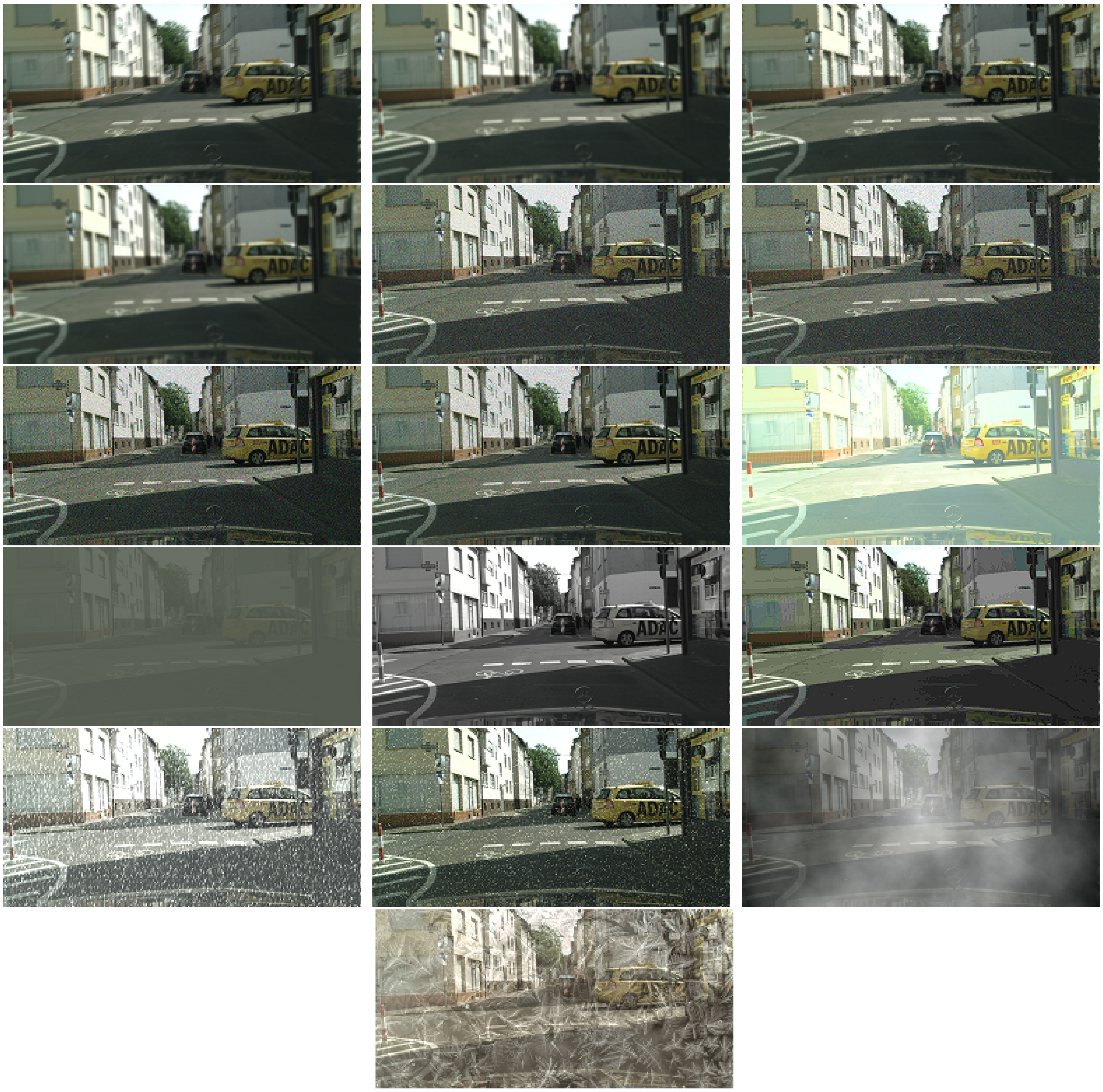


Figure A.1: Illustration of utilized image corruptions of ImageNet-C. First row: Motion blur, defocus blur, frosted glass blur. Second row: Gaussian blur, Gaussian noise, impulse noise. Third row: Shot noise, speckle noise, brightness. Fourth row: Contrast, saturate, JPEG. Fifth row: Snow, spatter, fog. Sixth row: frost.

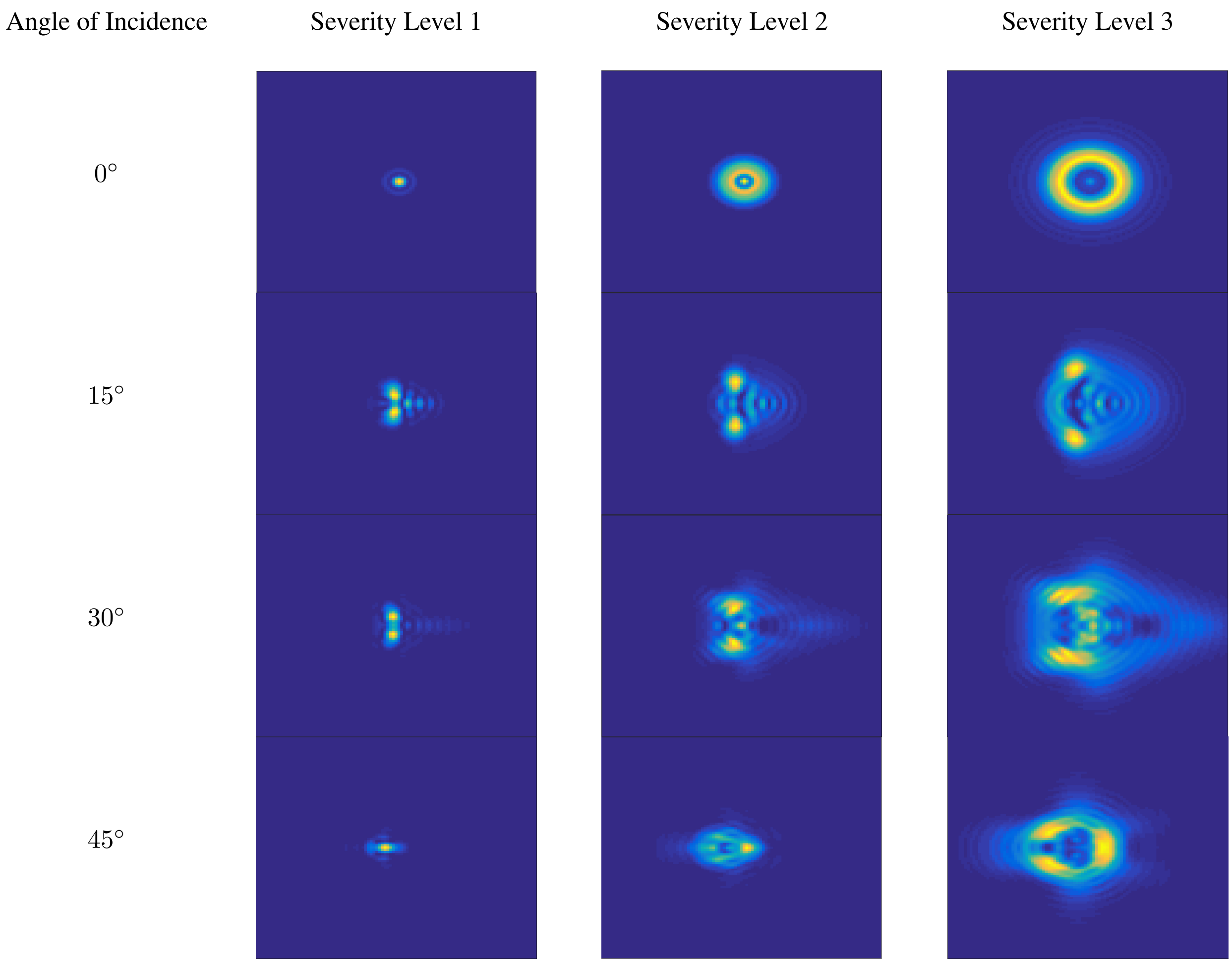


Figure A.2: The intensity distribution of used PSF kernels. The degree of the spatial distribution of intensity increases with the severity level. The shape of the PSF kernel depends on the image region, *i.e.*, the angle of incidence.

Figure A.3: Illustration of our proposed image corruptions. From left to right: Proposed noise model, PSF blur, and geometric distortion. Best viewed in color.

	Severity Level	Cityscapes	PASCAL VOC 2012	ADE20K
Gaussian Noise	1	13.2	18.6	18.3
	2	9.9	15.5	15.2
	3	6.8	12.4	12.1
	4	4.1	9.8	9.6
	5	1.7	7.3	7.3
Impulse Noise	1	11.2	16.7	16.5
	2	8.1	13.8	13.6
	3	6.4	12.1	12.0
	4	3.6	9.3	9.3
	5	1.6	7.2	7.4
Shot Noise	1	14.2	18.2	17.8
	2	10.5	14.9	14.3
	3	7.4	12.1	11.5
	4	3.9	8.9	8.2
	5	2.0	7.2	6.4
Speckle Noise	1	17.0	19.3	18.9
	2	14.5	17.1	16.6
	3	9.8	12.8	12.2
	4	7.9	11.0	10.3
	5	5.8	9.2	8.4
Intensity Noise	1	20.5	28.4	24.9
	2	18.6	22.1	23.1
	3	14.4	18.7	20.3
	4	10.8	15.1	15.5
	5	7.1	11.2	11.6

Table A.1: Averaged Signal-to-Noise ratios for corrupted variants of category image noise of the validation sets of Cityscapes, PASCAL VOC 2012, and ADE20K.

A.3. Signal-to-Noise Ratio for Image Corruptions

To make the severity levels of corruptions of category image noise mutually comparable, we provide in Tab. A.1 the SNR for this image corruption category.

B. Experiments

This section contains qualitative results and the remaining evaluation metric scores of the main paper. Based on DeepLabv3+, we evaluate the removal of atrous spatial pyramid pooling (ASPP), atrous convolutions (AC), and long-range link (LRL). We further replaced ASPP by Dense Prediction Cell (DPC) or applied global average pooling (GAP). Each ablated variant has been re-trained on the corresponding clean training data of Cityscapes, PASCAL VOC 2012, or ADE20K. To guarantee comparable results, we also re-trained the original, non-ablated models, though several publicly available checkpoints were available.

Besides DeepLabv3+, we have benchmarked several other semantic segmentation models as well.

B.1. Architectures

The DeepLabv3+ architecture functions[2] as reference model in our ablation study. We have benchmarked many other semantic segmentation models, such as FCN8s-VGG16 (trained by us), ICNet[3], DilatedNet[4], ResNet-38[5], PSPNet[6], and the recent Gated-ShapeCNN (GSCNN)[7], using mostly publicly available model checkpoints. Regarding PSPNet and ICNet, we conducted the benchmark using a PyTorch [4, 5] implementation[8].

B.2. Experimental Details

Hardware Setup. We have trained the models on machines equipped with four GTX 1080 Ti, each having 11 GB of memory or on a machine with two Titan RTX, each having 24 GB of memory.

Training Details. We set the crop size in every training for every dataset to 513, used a batch size of 16, as we always fine-tuned the batch normalization parameters. We applied the original training protocol of the developers of DeepLabv3+. We applied a polynomial learning rate with an initial learning rate of 0.007 or 0.01.

We re-trained 102 models for this benchmark: On ADE20K and the Cityscapes dataset, we re-trained 36 models each (six architectural ablations per network backbone; six network backbones in total). On PASCAL VOC, we re-trained 30 models, as we have evaluated on one network backbone less than on Cityscapes and ADE20K.

B.3. Evaluation Metrics

In the main paper, we predominantly use the Corruption Degradation (CD) to rate model robustness with respect to image corruptions, since the CD rates model robustness in terms of absolute performance. The relative Corruption Degradation (rCD), on the other hand, incorporates the respective model performance on clean data. The degradation on clean data is for both models (i.e., the model for which the robustness is to be rated, and the reference model) subtracted, resulting hence in a measure that gives a ratio of the absolute performance decrease in the presence of image corruption.

[2] https://github.com/tensorflow/models/tree/master/research/deeplab
[3] https://github.com/hszhao/ICNet
[4] https://github.com/fyu/dilation
[5] https://github.com/itijyou/ademxapp
[6] https://github.com/hszhao/PSPNet
[7] https://github.com/nv-tlabs/GSCNN
[8] https://github.com/meetshah1995/pytorch-semseg

B.4. Discussion of Architectural Properties

As mentioned in the main paper, we now discuss possible causes of the robustness of architectural design choices with respect to image corruptions in more detail.

Dense Prediction Cell. As mentioned in the main paper, a model with Dense Prediction Cell (DPC) might learn less multi-scale representations than a model with the Atrous Spatial Pyramid Pooling (ASPP) module. Whereas ASPP processes its input in parallel by three atrous convolution (AC) layers with large symmetric rates (6, 12, 18), DPC firstly processes the input by a single AC layer with small rate (1×6) [1, Fig. 5]. We hypothesize that DPC might learn less multi-scale representations than ASPP, which may be useful for common image corruptions (e.g., [2] shows that classification models are more robust to common corruption if the shape bias of a model is increased). When we test DPC on corrupted data, it cannot hence apply the same beneficial multi-scale cues (due to the comparable small atrous convolution with rate 1×6) as ASPP and may, therefore, perform worse.

Global Average Pooling. Global average pooling (GAP) increases performance on clean data on PASCAL VOC 2012, but not on the Cityscapes dataset or ADE20K. GAP averages 2048 activations of size 33×33 for our utilized training parameters. A possible explanation for the effectiveness of GAP on PASCAL VOC 2012 might be, that the Cityscapes dataset and ADE20K consist of both a notably larger number and spatial distribution of instances per image. Using GAP on these datasets might therefore not aid performance since important features may be lost due to averaging.

Long-Range Link. Removing the Long-Range Link (LRL) discards early representations. The degree of, *e.g.*, image noise is more pronounced on early CNN levels. Removing LRL tends hence to increase the robustness for a more shallow backbone as Xception-41 on PASCAL VOC 2012 and Cityscapes, as less corrupted features are linked from encoder to decoder. For a deeper backbone like ResNet-101, this behavior cannot be observed.

B.5. Qualitative Results

We provide qualitative results in this subsection. As mentioned in the main paper, blurred images cause the models to miss-classify pixels of classes covering small image regions, especially when far away. Please see Figure B.1 for an example. (a) A blurred validation image of the Cityscapes dataset and the corresponding ground truth in (b). (c) The prediction on the clean image overlaid with the ground truth (b). In this visualization, true-positives are alpha-blended, and false-positives, as well as false-negatives, remain unchanged. Hence, wrongly classified pixels can be easier identified. (d) The prediction on the blurred image overlaid with the ground truth (b). Whereas the *riders* and *persons* are mostly correctly classified in (c), *riders* in (d) are miss-classified as *persons*, and extensive areas of *road* are miss-classified as *sidewalk*. We used the reference model along with Xception-71 as network backbone to produce these predictions.

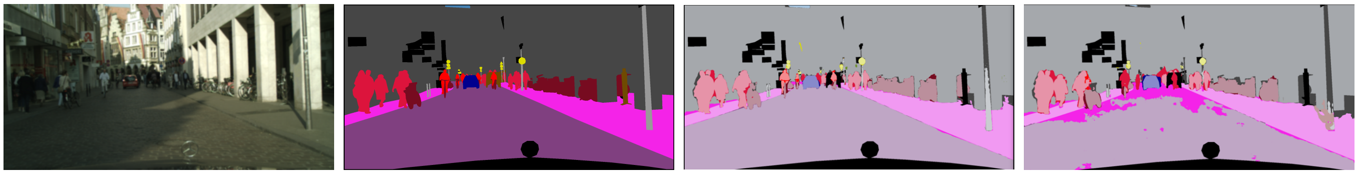
We report in the main paper, that–of the utilized image corruptions in this work–image noise affects model performance the most, as pointed out using mIoU scores. To give a visual example, we selected two noisy variants of a validation image of the Cityscapes dataset and show the predictions of the reference architecture using Xception-71 as network backbone in Figure B.2. The mIoU for both predictions is less than $15\,\%$.

Finally, we show qualitative results of every ablated architecture for one image corruption of category blur, noise, digital, and weather. Figure B.3 shows a blurred validation image of the Cityscapes dataset and the corresponding predictions. Note that the ablated variants w/o AC and w/ DPC are especially vulnerable. Figure B.4 shows a noisy validation image of the Cityscapes dataset. Note that the ablated variants w/o AC, w/o ASPP and w/ DPC are especially vulnerable. Figure B.5 and Figure B.6 show a validation image of PASCAL VOC 2012, corrupted by *brightness* and *snow*, respectively.

B.6. Experimental Results on Cityscapes

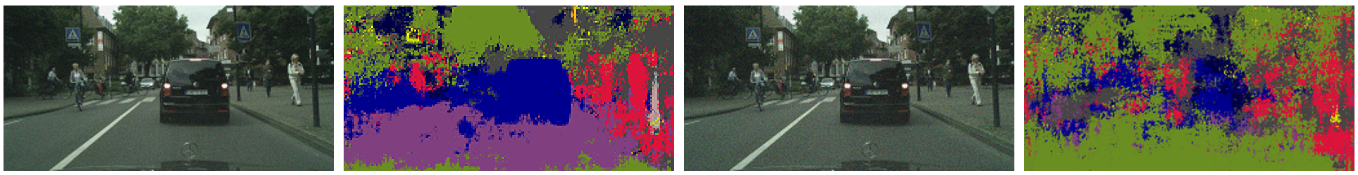
In this section, we provide a more detailed analysis of both the non-Deeplabv3+ and Deeplabv3+ based segmentation models. Figure B.7 illustrates the CD and rCD averaged for the proposed image corruption categories. Please note that the CD of image corruption “jpeg compression” of category digital is not included in this barplot. Contrary to the remaining image corruptions of that category, the respective CDs are considerably high (see Tab. B.1). FCN8s-VGG16 and DilatedNet are vulnerable to blur. The CD of *defocus blur* $124\,\%$ and $115\,\%$, respectively. However, DilatedNet is more robust against corruptions of category noise, digital, and weather than ICNet. For example, the CD of *intensity noise* is $92\,\%$. ResNet-38 is robust against corruptions of category weather. Both the CD and rCD for *fog* are roughly $65\,\%$. The CD of PSPNet is oftentimes less than $100\,\%$ (see Table B.1). GSCNN performs very well on digital corruptions (except *JPEG compression*) and weather corruptions, especially *fog* (CD is $44\,\%$, rCD is $34\,\%$.). The model is, however, vulnerable to image noise, as CD and rCD are always higher than $100\,\%$. We list the individual CD and rCD scores in Table B.1. Please find the absolute mIoU values in Table 1 (bottom) in the main paper.

Table B.2 contains the mIoU for clean and corrupted variants of the validation set of the Cityscapes dataset for several network backbones of the DeepLabv3+ architecture



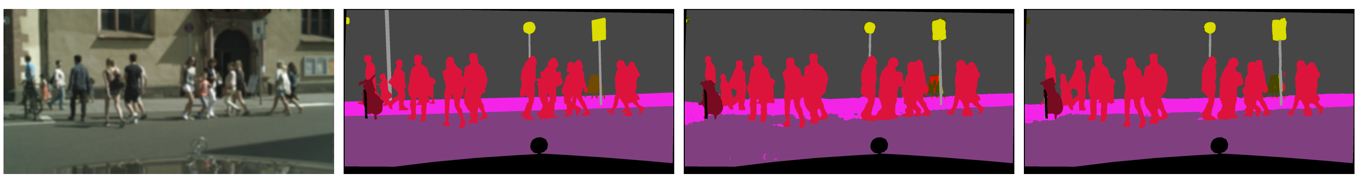
(a) Blurred validation image (b) ground truth (gt) (c) Overlay clean estimate + gt (d) Overlay blurred estimate + gt

Figure B.1: Prediction of the reference architecture (*i.e*. original DeepLabv3+) on blurred input, using Xception-71 as network backbone. (a) A blurred validation image of the Cityscapes dataset and corresponding ground truth (b). (c) Prediction on the clean image overlaid with the ground truth. True-positives are alpha-blended, false-positives and false-negatives remain unchanged. Hence, wrongly classified pixels can be easier spotted. (d) Prediction on the blurred image overlaid with the ground truth (b). Whereas the *riders* are mostly correctly classified in (c), they are in (d) miss-classified as *person*. Extensive areas of *road* are miss-classified as *sidewalk*.

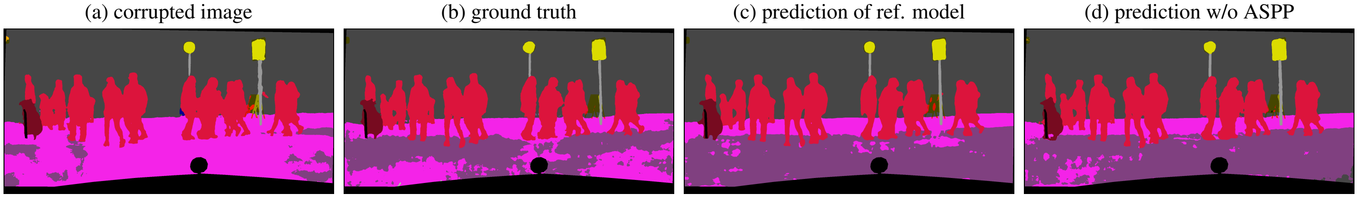


(a) Corrupted validation image (b) Prediction on (a) (c) Corrupted validation image (d) Prediction on (c)

Figure B.2: Drastic influence of image noise on model performance. (a) A validation image of Cityscapes is corrupted by the second severity level of Gaussian noise and respective prediction (b). (c) A validation image of Cityscapes is corrupted by the third severity level of Gaussian Noise and respective prediction (d). Predictions are produced by the reference model, using Xception-71 as the backbone.

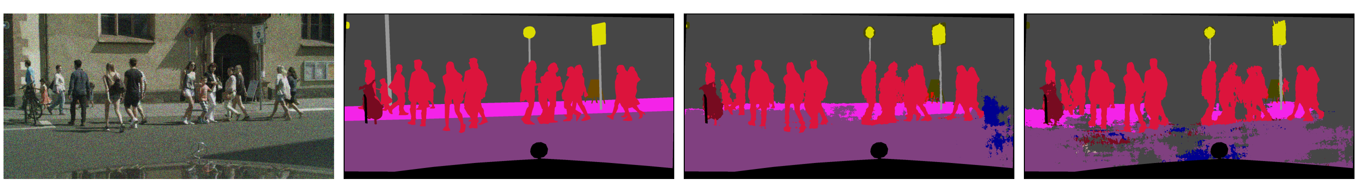


(a) corrupted image (b) ground truth (c) prediction of ref. model (d) prediction w/o ASPP

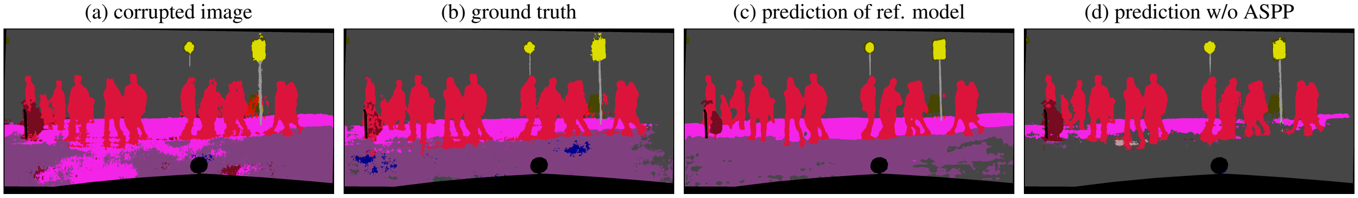


(e) prediction w/o AC (f) prediction w/ DPC (g) prediction w/o LRL (h) prediction w/ GAP

Figure B.3: Predictions of reference architecture and ablations on a blurred image. The ablated variants w/o AC, and w/ DPC are especially vulnerable to blur.

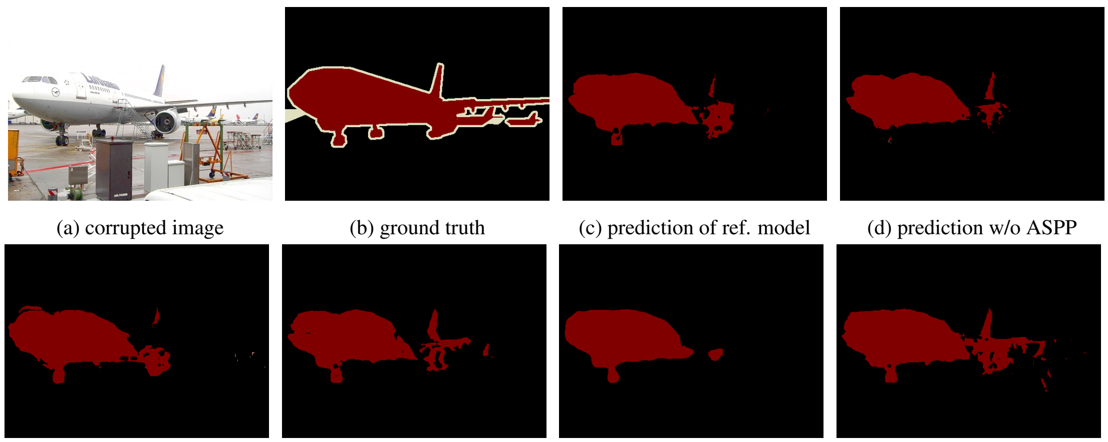


(a) corrupted image (b) ground truth (c) prediction of ref. model (d) prediction w/o ASPP



(e) prediction w/o AC (f) prediction w/ DPC (g) prediction w/o LRL (h) prediction w/ GAP

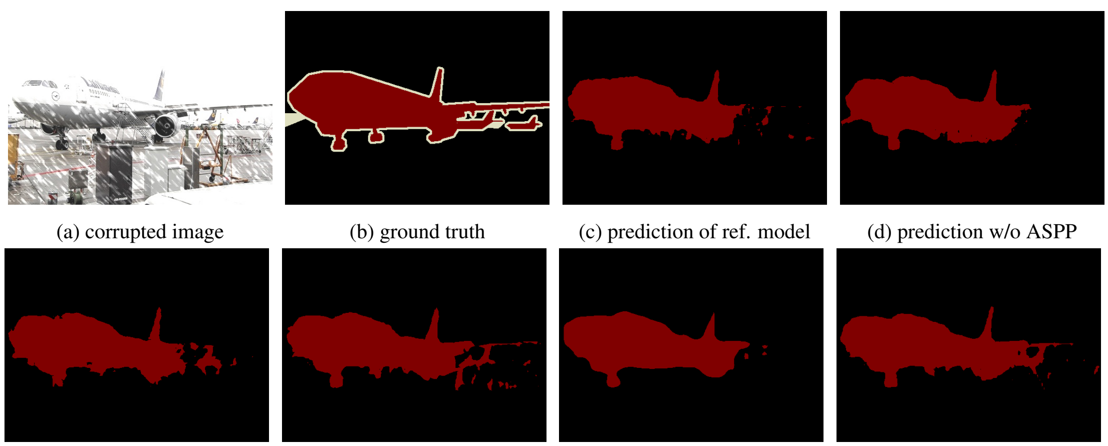
Figure B.4: Predictions of reference architecture and ablations on a noisy image. The ablated variants w/o AC, ASPP, and w/ DPC, GAP are especially vulnerable.



(a) corrupted image (b) ground truth (c) prediction of ref. model (d) prediction w/o ASPP

(e) prediction w/o AC (f) prediction w/ DPC (g) prediction w/o LRL (h) prediction w/ GAP

Figure B.5: Predictions of reference architecture and ablations on a validation image of PASCAL VOC 2012, corrupted by *brightness*.



(a) corrupted image (b) ground truth (c) prediction of ref. model (d) prediction w/o ASPP

(e) prediction w/o AC (f) prediction w/ DPC (g) prediction w/o LRL (h) prediction w/ GAP

Figure B.6: Predictions of the reference architecture and ablations on a validation image of PASCAL VOC 2012, corrupted by *snow*.

and each respective architectural modification. In addition to the main paper, where we discuss the CD score (see Figure 5 in the main paper), Figure B.9 illustrates the rCD (see equation 2 in the main paper) for each ablated variant evaluated on the Cityscapes dataset. In the following, we briefly discuss the ablated variants w.r.t. rCD.

Effect of ASPP. The rCD score is especially pronounced for geometrically distorted image data (146 % for Xception-41, 46 % for MobileNet-V2). The mIoU of Xception-41 on geometrically distorted data is low, resulting hence in a high CD and rCD. Regarding MobileNet-V2, the averaged mIoU is even similar to the other ablated variants (see the last column of Table B.2).

Effect of AC. As discussed in the main paper, AC show often an aiding effect against corruptions of type blur, noise, and geometric distortion (especially for ResNets and Xception-71). The rCD mostly has a similar tendency as the CD.

Effect of DPC. The rCD scores for the ablated variant without ASPP and with Dense Prediction Cell, generally show the same tendency as the CD illustrated in the main paper in Figure 5.

Effect of LRL. As mentioned in the main paper, this ablated variant is vulnerable to *intensity noise* (*defocus blur*), when applied in ResNet-101 (MobileNet-V2), since its rCD is 124 % (118 %). The rCD with respect to geometric distortion is especially high for MobileNet-V2 (155 %).

Effect of GAP. The rCD of Xception-71 and ResNet-101, using GAP, show a similar tendency as the CD discussed in the main paper. The rCD is low for Xception-71 w.r.t. geometric distortion.

Finally, we list for completeness the individual CD and rCD scores, evaluated on Cityscapes, in Table B.3 and Table B.4.

B.7. Experimental Results on PASCAL VOC 2012

Table B.5 contains the mIoU for clean and corrupted variants of the validation set of PASCAL VOC 2012 for several network backbones of the DeepLabv3+ architecture. In contrast to the model performance evaluated on Cityscapes, corruptions of category noise and weather have a less corrupting impact. Each backbone performs best on clean data when GAP is used. Each backbone performs significantly worse without ASPP. When GAP is used with ResNets and Xception-65, the resulting model is the best performing model on most types of image corruptions. Regarding Xception-41 and Xception-71, the ablated variant without LRL often has the highest mIoU w.r.t. image corruptions of category noise. Figure B.10 and Figure B.11 illustrates the CD and rCD for each ablated variant evaluated on PASCAL VOC 2012. In the following, we will briefly discuss the ablated variants w.r.t. rCD.

Effect of ASPP. For geometric distortion, the rCD of this ablated variant often shows except for Xception-41 a similar tendency the rCD as on Cityscapes (see Figure B.9). The rCD for ranges from 24 % (ResNet-101) to 62 % (Xception-65).

Effect of AC. As mentioned in the main paper, AC show no positive effect against blur. We explain this with the fundamentally different datasets. On Cityscapes, a model without AC often overlooks classes covering small image-regions, especially when far away. Such images are hardly present in PASCAL VOC 2012. An example of Cityscapes is illustrated in Figure B.1. The tendencies of CD and rCD are often similar to geometric distortion. As on Cityscapes, AC aids the robustness for ResNet-101 against this image corruption, but not for ResNet-50.

Effect of DPC. The harming effect of DPC with respect to image corruptions is especially present for Xception-71. As mentioned in the main paper, a possible explanation might be that the neural-architecture-search has been performed on Xception-71.

Effect of LRL. The rCD for this ablated variant is especially high w.r.t. geometrically distorted image data when applied in Xception-41 and ResNet-50 (144 % and 127 %, respectively). As the CD reported in the main paper, also the rCD of Xception-71 and Xception-41 for image noise is below 100 %.

Effect of GAP. Unlike the tendency of CD, the rCD of this ablated variant is rarely below 100.0 %. The rCD w.r.t. is high for geometric distortion.

Finally, we list the individual CD and rCD scores, evaluated on PASCAL VOC 2012, in Table B.6 and Table B.7.

B.8. Experimental Results on ADE20K

Table B.8 contains the mIoU for clean and corrupted variants of the validation set of ADE20K for several network backbones of the DeepLabv3+ architecture. When comparing the respective reference model of each backbone (*i.e.* no ablated variants), Xception-71 performs best for every type of image corruption. MobileNet-V2 (ResNets) oftentimes performs best when DPC (GAP) is used. Xception-41 and Xception-71 perform best on clean data when DPC is used. Most backbones without ASPP perform significantly worse than respective reference. Figure B.12 and Figure B.13 illustrates the CD and rCD for each ablated variant evaluated on ADE20K. As mentioned in the main paper, the CD is–except for models without ASPP–oftentimes around 100 %. In the following, we will briefly discuss the ablated variants w.r.t. CD and rCD.

Effect of ASPP. The rCD is in general above 100 % for both Xception-65 and Xception-71, and below 100 % for remaining backbones. The performance gap w.r.t. the reference model is for the aforementioned Xception-based back-

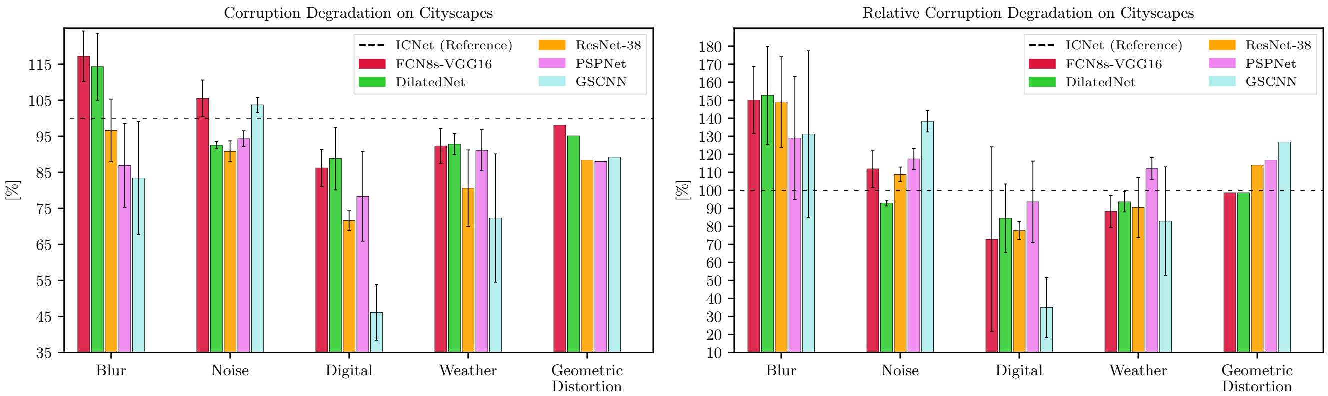


Figure B.7: CD (left) and rCD (right) evaluated on Cityscapes for ICNet (set as reference architecture), FCN8s-VGG16, DilatedNet, ResNet-38, PSPNet, GSCNN w.r.t. image corruptions of category blur, noise, digital, weather, and geometric distortion. Each bar except for geometric distortion is averaged within a corruption category (error bars indicate the standard deviation). The CD of image corruption "jpeg compression" of category digital is not included in this barplot, since, contrary to the remaining image corruptions of that category, the respective CDs range between 107 % and 133 %. Bars above 100 % represent a decrease in performance compared to the reference architecture. FCN8s-VGG16 and DilatedNet are vulnerable to corruptions of category blur. DilatedNet is more robust against corruptions of category noise, digital, and weather than the reference. ResNet-38 is robust against corruptions of category weather. The rCD of PSPNet is oftentimes higher than 100 % for each category. GSCNN is vulnerable to image noise. The rCD is considerably high, indicating a high decrease of mIoU in the presence of this corruption. Best viewed in color.

bones significantly less than for the remaining backbones.

Effect of AC. The removal of AC decreases the performance slightly for most backbones against corruptions of category digital and weather.

Effect of DPC. As on PASCAL VOC 2012 and Cityscapes, applying DPC oftentimes decreases the robustness, especially for Xception-71 against most image corruptions. As on Cityscapes, using DPC along Xception-71, results in the best-performing model on clean data.

Effect of LRL. The removal of LRL impacts especially Xception-71 against image noise.

Effect of GAP. When GAP is applied, the models perform generally most robust.

Finally, we list the individual CD and rCD scores, evaluated on ADE20K, in Table B.9 and Table B.10.

B.9. Performance without ASPP

As mentioned in the main paper, we provide a more detailed evaluation for the ablated architecture without ASPP for every dataset, in this subsection. The Atrous Spatial Pyramid Pooling (ASPP) module reduces, in general, the model performance significantly. On PASCAL VOC 2012, the mIoU on clean data reduces between 5.9 % (Xception-65) and 12.0 % (ResNet-50). On ADE20K, the mIoU decreases between 1.2 % (Xception-65) and 7.7 % (ResNet-50). On Cityscapes, the mIoU decreases between 2.4 % (Xception-41) and 7.1 % (MobileNet-V2). Therefore, the corresponding CD scores are oftentimes considerably high. On PASCAL VOC 2012 and ADE20K, for example, the CD score for the ablated variant w/o ASPP is the highest w.r.t. every image corruption and for every network backbone (see bold values Table B.6 and Table B.9). Regarding the evaluation on Cityscapes (see Table B.3), the CD score of the ablated variant w/o ASPP is often high against image corruptions of category blur for ResNets and MobileNet-V2. Stronger backbones, on the other hand, as Xception-based ones, perform better without ASPP and have thus a lower CD.

B.10. Performance with respect to Individual Severity Levels

We illustrate in Fig. B.8 the model performance evaluated on every dataset with respect to individual severity levels. The Figure shows the degrading performance with increasing severity level for some candidates of category blur, noise, digital, and weather of a reference model and all corresponding architectural ablations. Please see the caption for discussion.

References

[1] Liang-Chieh Chen, Maxwell D. Collins, Yukun Zhu, George Papandreou, Barret Zoph, Florian Schroff, Hartwig Adam, and Jonathon Shlens. Searching for Efficient Multi-Scale Architectures for Dense Image Prediction. In *NIPS*, 2018.

[2] R. Geirhos, P. Rubisch, C. Michaelis, M. Bethge, F. A. Wichmann, and W. Brendel. ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. In *ICLR*, May 2019.

[3] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

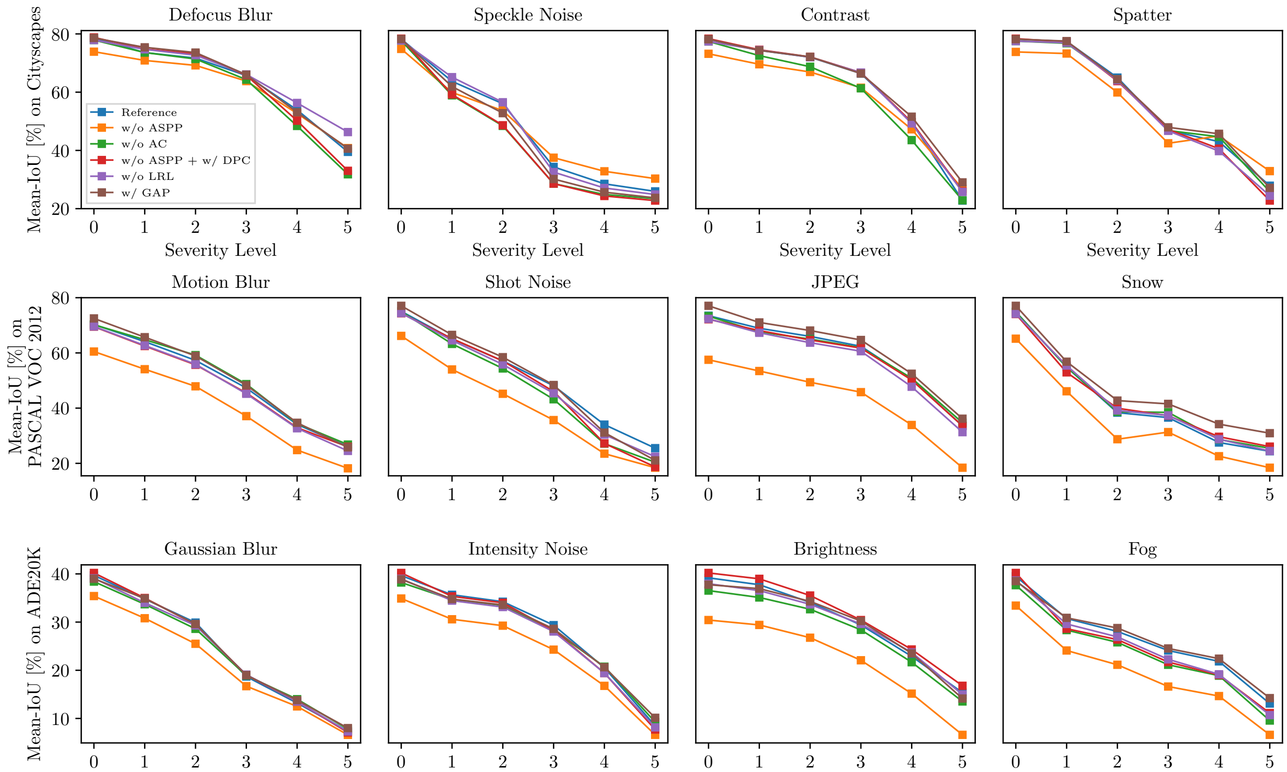


Figure B.8: Model performance (mIoU) for many candidates with respect to the image corruption categories blur (first column), noise (second column), digital (third column), and weather (fourth column) for a reference model and all corresponding architectural ablated variants, evaluated for every severity levels on Cityscapes, PASCAL VOC 2012, and ADE20K. Severity level 0 corresponds to clean data. **First row:** Xception-71 evaluated on the Cityscapes dataset for defocus blur, speckle noise, contrast, and spatter. **Second row:** ResNet-101 evaluated on PASCAL VOC 2012 for motion blur, shot noise, JPEG, and snow. **Third row:** Xception-41 evaluated on ADE20K for Gaussian blur, intensity noise, brightness, and fog. The ablated variant without ASPP oftentimes has the lowest mIoU. However, it performs best on speckle noise for severity level 3 and above. The mIoU of the ablated variant without AC is relatively low for defocus blur and contrast. The mIoU of the ablated variant without ASPP and with DPC is relatively low for speckle noise, shot noise (for severity level 4 and 5), spatter. The mIoU of the ablated variant without LRL is relatively high for speckle noise and shot noise. The mIoU of the ablated variant with GAP is high for PASCAL VOC 2012 on clean data and low for speckle noise.

	Blur					**Noise**					**Digital**				**Weather**				
Deeplab-v3+ Backbone	Motion	Defocus	Frosted Glass	Gaussian	PSF	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
ICNet	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0	**100.0**	100.0	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
FCN8s-VGG16	**105.6**	**124.3**	119.6	**119.1**	110.8	101.8	103.0	103.2	104.1	**115.6**	79.2	91.2	88.2	119.4	94.7	98.4	85.8	90.2	98.1
DilatedNet	102.6	115.1	**128.3**	111.4	**111.8**	92.2	93.9	91.3	93.3	91.9	80.2	**100.7**	85.4	107.3	93.4	97.3	89.8	90.7	95.1
ResNet-38	83.7	99.2	107.8	95.5	72.0	94.3	91.8	91.5	85.5	91.2	67.8	73.8	73.3	129.2	92.4	77.9	64.7	87.4	88.4
PSPNet	74.1	84.6	105.7	83.3	66.3	97.1	92.4	94.7	91.1	96.2	67.1	72.1	95.7	119.1	97.8	82.5	90.2	94.1	88.0
GSCNN	75.9	75.1	110.4	72.2	56.5	**103.2**	**106.4**	**104.3**	**104.4**	100.0	40.9	57.0	40.4	**133.2**	93.5	75.8	44.1	75.7	89.2
ICNet	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
FCN8s-VGG16	119.1	**167.3**	160.1	**153.8**	779.8	104.3	106.1	106.6	109.9	132.5	54.0	84.6	79.9	142.7	93.1	99.1	75.3	85.5	98.6
DilatedNet	**120.5**	152.2	195.4	142.8	**1117.9**	92.3	95.0	90.8	94.4	91.8	64.0	**109.9**	79.5	123.8	94.3	102.5	87.8	89.9	98.6
ResNet-38	114.2	152.9	185.3	143.5	388.9	111.2	107.2	107.4	103.1	115.2	70.6	82.1	80.0	197.2	107.7	89.5	63.7	100.8	114.0
PSPNet	93.7	120.0	185.3	116.8	256.0	117.7	110.2	114.6	116.8	127.9	73.5	82.1	**125.2**	179.7	**117.9**	**101.7**	**114.7**	**113.7**	116.8
GSCNN	109.7	105.9	**211.0**	98.3	85.1	**131.2**	**136.4**	**134.2**	**147.9**	**141.6**	20.2	58.1	26.5	**216.0**	115.0	95.0	33.7	87.9	**126.8**

Table B.1: CD (top) and rCD (bottom) for corrupted variants of the validation set of the Cityscapes dataset for several non-Deeplabv3+ based architectures. ICNet is used as reference model. Highest CD and rCD per corruption is bold.

[4] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic Differentiation in PyTorch. In *NIPS Autodiff Workshop*, 2017.

[5] Meet P. Shah. Semantic Segmentation Architectures Implemented in PyTorch. *https://github.com/meetshah1995/pytorch-semseg*, 2017.

		Blur					Noise					Digital				Weather				
Deeplab-v3+ Backbone	Clean	Motion	Defocus	Frosted Glass	Gaussian	PSF	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
MobileNet-V2	**72.0**	**53.5**	49.0	**45.3**	49.1	**70.5**	6.4	7.0	6.6	16.6	26.9	51.7	46.7	32.4	27.2	**13.7**	38.9	**47.4**	**17.3**	65.5
w/o ASPP	64.9	45.5	40.4	39.0	41.5	63.3	7.7	8.7	8.9	19.9	28.4	41.7	36.1	27.6	20.7	13.0	36.8	37.8	14.0	61.9
w/o AC	71.2	52.1	49.1	42.8	49.3	69.8	3.6	7.2	4.5	19.6	29.2	49.8	46.2	31.4	**28.1**	10.0	**44.6**	45.2	16.7	63.5
w/ DPC	71.6	49.4	42.2	43.7	43.8	69.2	3.5	4.9	3.9	16.1	27.4	45.0	38.6	30.1	24.1	9.8	42.8	43.9	14.0	62.2
w/o LRL	71.1	49.7	43.9	44.4	45.1	68.9	1.9	2.6	2.5	19.6	26.6	49.1	43.5	32.2	26.3	10.4	39.5	44.9	14.7	60.9
w/ GAP	71.4	52.2	**50.6**	43.3	**51.8**	69.8	**8.5**	**10.9**	**10.8**	**26.0**	**32.5**	**51.8**	**47.3**	**35.3**	25.7	12.7	43.4	45.1	12.6	**66.0**
ResNet-50	76.6	**58.5**	**56.6**	47.2	**57.7**	74.8	6.5	7.2	10.0	31.1	30.9	**58.2**	**54.7**	41.3	27.4	**12.0**	**42.0**	55.9	**22.8**	**69.5**
w/o ASPP	71.4	52.3	50.7	41.2	52.0	69.7	10.1	11.3	**13.8**	31.2	33.3	50.2	48.4	37.0	25.3	12.0	38.6	42.7	18.7	65.9
w/o AC	76.0	56.7	53.1	**47.3**	54.1	73.8	2.4	6.1	5.1	25.5	25.7	56.8	51.4	38.9	27.6	9.7	40.8	52.0	20.1	66.9
w/ DPC	**76.9**	57.0	54.7	46.9	56.2	74.2	**10.7**	**12.6**	13.6	**33.1**	32.0	54.5	53.6	**41.5**	25.1	11.4	41.3	**56.3**	20.4	68.6
w/o LRL	75.6	57.9	54.4	46.4	55.6	73.8	7.9	9.3	11.2	31.8	**34.7**	56.2	51.6	40.2	**28.5**	11.9	41.4	55.4	21.1	67.9
w/ GAP	76.5	56.7	55.7	45.8	57.4	**75.2**	5.5	7.8	9.5	31.3	34.5	57.7	51.4	41.1	28.3	10.5	40.4	54.5	20.1	68.5
ResNet-101	77.1	**59.1**	56.3	47.7	57.3	75.2	**13.2**	**13.9**	**16.3**	**36.9**	**39.9**	**59.2**	54.5	**41.5**	**37.4**	**11.9**	**47.8**	55.1	**22.7**	**69.7**
w/o ASPP	71.1	53.8	50.6	42.2	51.7	68.8	9.5	9.8	12.7	30.7	32.5	52.1	48.3	36.7	33.2	13.3	43.5	47.8	23.2	66.4
w/o AC	75.7	57.9	52.5	46.6	53.9	73.3	8.4	11.0	11.6	31.5	28.8	53.5	53.1	39.1	34.2	9.9	44.7	55.0	20.0	65.5
w/ DPC	77.0	58.5	53.5	46.7	54.8	75.3	11.7	12.1	15.6	36.4	35.5	53.7	54.3	39.8	30.9	10.1	44.0	**56.0**	19.3	68.6
w/o LRL	76.5	58.7	54.6	47.5	55.7	74.3	9.1	8.3	12.1	33.5	30.3	57.0	**57.6**	40.9	35.7	9.3	44.3	55.4	20.8	69.2
w/ GAP	**77.3**	58.7	**56.9**	**48.4**	**57.8**	**75.9**	8.2	7.4	11.6	32.0	32.8	55.6	55.8	39.3	36.4	11.5	44.8	52.5	22.6	69.0
Xception-41	**77.8**	61.6	54.9	51.0	**54.7**	**76.1**	17.0	17.3	21.6	43.7	**48.6**	63.6	56.9	**51.7**	**38.5**	18.2	46.6	57.6	20.6	**73.0**
w/o ASPP	75.4	59.7	55.5	47.4	55.4	73.1	15.1	14.4	19.7	40.7	43.6	60.4	52.5	46.8	37.0	18.0	47.2	52.4	22.1	68.4
w/o AC	77.4	62.2	**55.6**	51.3	54.5	75.4	17.7	15.7	22.1	42.8	46.5	61.6	54.9	47.8	34.3	17.8	46.6	**59.1**	20.9	70.9
w/ DPC	77.5	60.6	53.0	50.8	52.5	75.8	15.1	10.7	20.3	42.7	48.4	63.6	53.4	46.0	36.0	17.6	**50.0**	56.7	20.6	71.8
w/o LRL	76.8	**62.3**	53.2	50.6	53.0	75.1	**21.3**	**19.2**	**27.6**	**49.3**	51.7	63.9	55.2	48.0	33.8	20.5	48.3	57.6	23.9	70.8
w/ GAP	77.1	61.5	54.8	**53.1**	53.9	75.6	20.0	16.4	24.8	43.4	46.6	**65.7**	**57.6**	50.4	36.2	16.5	48.6	56.8	22.6	72.5
Xception-65	**78.4**	**63.9**	59.1	52.8	**59.2**	**76.8**	15.0	10.6	19.8	42.4	46.5	**65.9**	**59.1**	46.1	31.4	**19.3**	50.7	**63.6**	**23.8**	**72.7**
w/o ASPP	75.8	61.6	56.1	51.8	54.6	74.1	14.3	7.7	18.8	39.0	41.6	62.0	57.2	43.1	29.7	15.6	46.9	60.3	23.4	70.6
w/o AC	77.7	63.9	58.7	51.5	57.8	75.7	14.1	14.8	19.5	41.9	45.1	63.9	58.3	42.9	35.0	15.7	**51.4**	60.9	21.4	71.8
w/ DPC	77.7	62.4	55.0	50.4	54.5	74.7	8.9	4.8	13.2	37.1	**47.7**	62.5	48.4	45.4	30.3	17.3	47.1	59.6	21.9	70.7
w/o LRL	77.7	64.5	58.6	49.5	57.9	75.9	15.1	12.0	19.9	42.1	45.9	63.8	57.9	46.1	**35.9**	18.4	46.3	63.5	22.0	71.4
w/ GAP	78.4	63.9	**59.4**	**53.5**	58.8	76.2	**18.8**	**15.4**	**23.7**	**43.7**	45.7	65.2	56.5	**48.0**	31.5	18.8	49.4	59.1	20.7	71.0

Table B.2: Mean IoU for clean and corrupted variants of the validation set of the Cityscapes dataset for several network backbones of the DeepLabv3+ architecture and respective architectural ablations. Every mIoU is averaged over all available severity levels, except for corruptions of category noise where only the first three severity levels are considered. The standard deviation for image corruptions of category noise is 0.2 or less. Highest mIoU per corruption is bold.

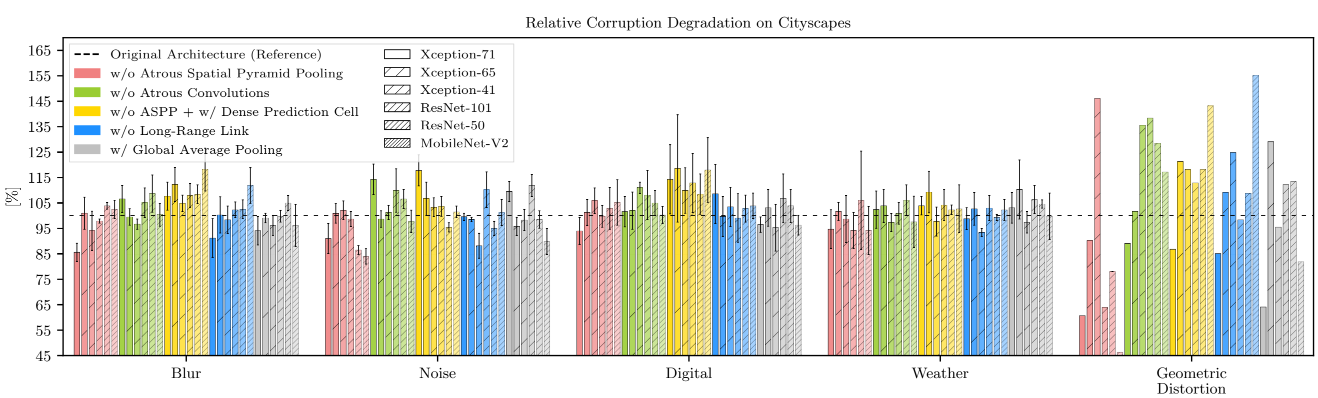


Figure B.9: Relative CD evaluated on Cityscapes for the proposed ablated variants of the DeepLabv3+ architecture w.r.t. image corruptions, employing six different network backbones. Each bar except for geometric distortion is averaged within a corruption category (error bars indicate the standard deviation). Bars above 100 % represent a relative decrease in performance compared to the respective reference architecture. Each ablated architecture is re-trained on the original training dataset. Removing ASPP may decrease performance significantly. The low rCD for geometric distortion indicates that the relative decrease of performance for this ablated variant is low. AC affect model performance, particularly against geometric distortion. The relative CD is often high against most image corruptions when DPC is used. The effect of GAP depends strongly on the network backbone. Best viewed in color.

	Blur					Noise					Digital				Weather				
Deeplab-v3+ Backbone	Motion	Defocus	Frosted Glass	Gaussian	PSF	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
MobileNet-V2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**117.2**	**116.8**	**111.5**	**115.0**	**124.2**	98.6	98.1	97.6	96.1	97.9	**120.6**	**120.0**	**107.1**	**109.0**	100.7	**103.5**	**118.2**	104.0	**110.4**
w/o AC	103.1	99.9	104.6	99.8	102.2	103.0	99.7	102.2	96.4	96.9	104.0	101.1	101.5	98.8	104.2	90.8	104.1	100.7	105.7
w/ DPC	108.9	113.4	102.9	110.4	104.2	103.1	102.3	102.9	**100.6**	99.3	113.9	115.3	103.4	104.3	**104.5**	93.7	106.7	103.9	109.4
w/o LRL	108.2	110.0	101.7	108.0	105.3	**104.8**	**104.8**	**104.4**	96.3	**100.0**	105.5	106.1	100.3	101.2	103.8	99.1	104.8	103.1	113.2
w/ GAP	102.8	96.9	103.6	94.8	102.2	97.7	95.8	95.5	88.7	92.3	99.9	98.9	95.7	102.1	101.1	92.8	104.3	**105.6**	98.4
ResNet-50	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**115.1**	**113.6**	**111.3**	**113.5**	**120.3**	96.2	95.5	95.8	99.9	96.4	**119.0**	**113.9**	**107.3**	102.9	100.1	**105.8**	**130.1**	**105.3**	**111.8**
w/o AC	104.4	108.1	99.7	108.5	104.2	**104.3**	**101.2**	**105.4**	**108.1**	**107.5**	103.2	107.4	104.1	99.8	**102.6**	102.0	109.0	103.5	108.6
w/ DPC	103.6	104.3	100.5	103.7	102.3	95.5	94.2	96.0	97.1	98.3	108.6	102.4	99.7	**103.3**	100.7	101.0	99.1	103.1	103.0
w/o LRL	101.5	104.9	101.5	105.1	104.2	98.5	97.7	98.6	99.1	94.4	104.6	106.8	101.8	98.5	100.1	101.0	101.3	102.2	105.1
w/ GAP	104.3	102.1	102.7	100.7	98.3	101.1	99.3	100.6	99.7	94.7	101.2	107.3	100.3	98.8	101.8	102.8	103.3	103.4	103.2
ResNet-101	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**113.2**	**113.1**	**110.5**	**113.3**	**125.6**	104.3	104.8	104.4	**109.9**	112.2	**117.6**	**113.4**	**108.3**	106.7	98.4	**108.1**	**116.2**	99.4	111.0
w/o AC	103.1	108.6	102.2	108.0	107.5	105.6	103.3	105.7	108.6	118.5	114.2	103.1	104.2	105.2	102.4	105.9	100.3	103.5	**113.9**
w/ DPC	101.5	106.5	101.9	105.8	99.8	101.8	102.1	100.8	100.8	107.2	113.5	100.4	103.0	**110.5**	102.1	107.2	98.0	**104.4**	103.5
w/o LRL	101.2	103.9	100.5	103.8	103.8	104.8	106.4	105.1	105.4	**115.9**	105.6	93.1	101.1	102.8	**103.1**	106.7	99.5	102.5	101.5
w/ GAP	101.1	98.6	98.7	98.9	97.1	**105.8**	**107.5**	**105.7**	107.8	111.8	108.9	97.1	103.7	101.7	100.5	105.6	105.8	100.2	102.2
Xception-41	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0
w/o ASPP	**105.1**	98.5	**107.5**	98.5	**112.2**	**102.4**	103.5	**102.4**	**105.3**	**109.8**	**108.9**	**110.2**	110.2	102.3	100.2	98.8	**112.4**	98.2	**116.9**
w/o AC	98.5	98.3	99.5	100.4	102.6	99.1	102.0	99.3	101.6	104.0	105.6	104.7	108.2	106.8	100.5	**100.0**	96.6	99.6	107.5
w/ DPC	102.8	**104.2**	100.5	**104.9**	101.0	102.4	**108.0**	101.6	101.9	100.4	100.0	108.1	**111.9**	104.1	100.8	93.6	102.1	100.0	104.4
w/o LRL	98.2	103.6	100.9	103.7	103.9	94.9	97.7	92.3	90.1	94.0	99.2	103.9	107.7	**107.6**	97.2	96.9	100.0	95.9	108.0
w/ GAP	100.3	100.2	95.7	101.7	101.8	96.5	101.1	95.9	100.5	103.9	94.3	98.5	102.6	103.8	**102.1**	96.2	101.9	97.5	101.6
Xception-65	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**106.4**	107.2	102.0	111.4	**111.7**	100.8	103.3	101.4	105.9	**109.3**	**111.2**	104.7	105.5	**102.5**	**104.5**	107.7	109.1	100.6	**107.5**
w/o AC	100.1	101.0	102.6	103.3	104.8	101.1	95.3	100.5	100.9	102.7	105.6	102.1	**106.0**	94.8	104.4	98.7	107.4	103.1	103.1
w/ DPC	104.0	**109.8**	105.1	**111.6**	108.9	**107.2**	**106.5**	**108.3**	**109.3**	97.9	109.9	**126.3**	101.4	101.7	102.4	107.4	111.0	102.5	107.1
w/o LRL	98.3	101.1	**107.0**	103.2	104.0	99.8	98.4	99.9	100.6	101.3	106.1	103.1	99.9	93.4	101.1	**108.9**	100.3	102.4	104.8
w/ GAP	100.0	99.1	98.5	101.0	102.5	95.5	94.6	95.2	97.8	101.5	101.9	106.6	96.5	99.9	100.6	102.6	**112.4**	**104.2**	106.3
Xception-71	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0
w/o ASPP	**109.6**	103.6	101.1	105.2	**115.1**	95.7	95.6	96.5	102.4	110.9	114.6	106.0	**108.3**	100.5	101.6	102.9	**116.1**	97.9	**105.9**
w/o AC	105.3	**107.6**	100.4	**105.7**	100.9	**108.4**	105.8	110.2	**114.3**	114.0	101.4	**107.4**	102.0	99.1	100.9	100.8	108.2	98.2	99.7
w/ DPC	103.7	103.7	98.8	105.6	97.9	108.9	**109.0**	**110.8**	113.9	**115.5**	110.2	98.2	103.4	**113.7**	**105.5**	**103.7**	100.2	99.2	96.0
w/o LRL	99.8	94.0	**102.8**	95.4	98.7	101.1	101.7	101.5	99.9	100.5	**111.0**	98.7	105.4	106.9	102.4	103.6	99.3	98.5	98.7
w/ GAP	99.8	97.8	91.8	99.3	94.0	106.1	102.8	106.9	107.3	109.1	100.1	96.2	97.7	99.3	102.6	98.9	105.4	99.0	90.8

Table B.3: CD for corrupted variants of the validation set of the Cityscapes dataset for several network backbones of the DeepLabv3+ architecture and respective architectural ablations. Highest CD per corruption is bold.

	Blur					Noise					Digital				Weather				
Deeplab-v3+ Backbone	Motion	Defocus	Frosted Glass	Gaussian	PSF	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
MobileNet-V2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0
w/o ASPP	104.8	106.3	96.8	102.1	101.3	87.2	86.3	85.7	81.2	80.8	114.0	114.0	94.1	98.7	88.9	84.9	109.9	93.0	46.3
w/o AC	103.2	96.1	**106.4**	95.8	89.0	103.0	98.3	101.9	93.1	93.1	105.5	99.0	100.5	96.3	104.9	80.5	105.4	99.6	117.2
w/ DPC	**119.9**	**127.8**	104.3	**121.3**	**152.5**	103.8	102.6	103.4	**100.2**	97.9	**130.9**	**130.5**	**104.7**	**106.0**	**105.9**	87.1	**112.5**	105.1	143.2
w/o LRL	115.5	118.1	100.0	113.8	141.6	**105.4**	**105.4**	**104.8**	92.8	97.9	108.5	109.2	98.1	100.0	104.1	95.5	106.5	103.0	**155.2**
w/ GAP	103.5	90.4	105.0	85.8	100.6	95.8	93.0	92.6	81.9	86.1	96.7	95.2	91.1	102.0	100.5	84.7	106.6	**107.3**	81.9
ResNet-50	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	106.1	103.7	102.8	102.9	98.2	87.5	86.6	86.5	88.4	83.3	**115.1**	105.2	97.5	93.8	92.1	94.9	**139.3**	98.0	78.0
w/o AC	106.7	**114.5**	97.5	**115.9**	126.3	**104.9**	**100.7**	**106.5**	**111.0**	**110.0**	103.9	112.5	**105.0**	98.4	102.6	101.7	116.3	103.9	**128.5**
w/ DPC	**110.3**	111.1	102.1	110.2	**154.5**	94.5	92.8	95.2	96.3	98.3	121.6	106.6	100.5	**105.5**	101.5	102.8	99.8	**105.1**	118.1
w/o LRL	98.2	106.0	99.5	106.5	106.5	96.7	95.5	96.8	96.5	89.5	105.4	109.8	100.3	95.8	98.7	99.0	98.3	101.4	108.8
w/ GAP	109.7	104.3	**104.7**	101.3	73.1	101.4	99.0	100.7	99.4	91.9	102.5	**114.9**	100.3	98.1	**102.3**	**104.5**	106.9	104.9	113.4
ResNet-101	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	96.5	98.6	98.3	98.3	119.0	96.4	97.0	96.1	100.6	103.6	106.5	**100.5**	96.7	95.5	88.7	94.0	105.9	88.1	63.9
w/o AC	99.3	111.4	99.2	110.4	**125.1**	105.5	102.3	105.5	110.1	**126.1**	124.6	100.1	103.0	104.7	101.1	105.8	94.3	102.4	**138.4**
w/ DPC	103.0	**113.2**	**103.1**	**112.2**	92.1	102.3	102.7	101.0	101.0	111.4	**130.4**	100.4	104.7	**116.3**	102.7	**112.6**	95.5	**106.1**	112.9
w/o LRL	99.5	105.5	98.9	105.4	120.0	105.7	107.9	106.1	107.1	124.1	109.5	83.7	100.3	102.9	**103.3**	110.0	96.3	102.4	98.4
w/ GAP	**103.8**	98.2	98.5	98.9	74.0	**108.3**	**110.6**	**108.3**	**112.8**	119.8	121.7	95.1	**106.8**	103.4	101.0	110.8	**112.9**	100.7	112.2
Xception-41	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0	**100.0**	100.0
w/o ASPP	97.7	86.9	**104.9**	86.8	**133.6**	99.4	100.9	99.1	101.9	**109.2**	106.3	109.9	109.9	97.7	96.4	90.5	**114.4**	93.4	**146.1**
w/o AC	94.3	95.3	97.8	99.4	116.9	98.3	102.2	98.5	101.6	106.0	**112.1**	108.2	114.0	**109.8**	100.1	98.9	91.2	98.9	135.6
w/ DPC	**104.7**	**106.9**	99.8	**108.2**	96.2	**102.7**	**110.4**	**101.7**	**102.2**	99.7	97.8	**115.3**	**120.9**	105.6	100.5	88.1	102.8	99.5	118.1
w/o LRL	89.6	102.9	98.0	103.0	97.2	91.4	95.3	87.5	80.8	86.0	91.0	103.3	110.5	109.4	94.5	91.5	95.2	92.6	124.8
w/ GAP	96.7	97.6	89.7	100.5	88.3	94.1	100.4	93.1	99.0	104.6	80.9	93.8	102.4	104.3	**101.8**	91.4	100.9	95.4	95.5
Xception-65	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	98.0	101.7	93.6	**110.6**	106.5	97.0	100.4	97.4	102.1	**107.4**	109.7	96.4	101.1	98.1	101.7	104.3	104.8	96.1	90.2
w/o AC	95.0	98.1	101.8	103.1	121.2	100.2	92.7	99.3	99.3	102.2	109.3	100.5	**107.7**	90.7	**104.7**	94.9	113.0	103.0	101.7
w/ DPC	**105.0**	**117.0**	106.5	120.8	**181.5**	**108.5**	**107.4**	**110.2**	**112.8**	94.2	**121.2**	**152.0**	100.0	**100.9**	102.1	110.6	122.2	102.2	121.3
w/o LRL	90.4	98.3	**109.8**	102.7	108.8	98.6	96.8	98.5	98.8	99.7	110.2	102.5	97.5	88.8	100.1	**113.1**	95.5	101.9	109.2
w/ GAP	99.7	97.8	97.0	102.0	132.6	93.9	92.8	93.4	96.4	102.4	104.8	113.8	94.1	99.7	100.7	104.4	**130.2**	**105.7**	**129.1**
Xception-71	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**
w/o ASPP	91.2	81.3	84.3	85.5	49.1	86.9	87.3	87.3	91.1	102.4	99.6	88.8	99.0	88.4	94.2	88.3	107.3	89.1	60.7
w/o AC	107.9	**112.4**	97.9	108.1	76.3	110.0	106.5	112.6	120.5	121.8	97.1	**111.5**	101.0	96.7	99.9	98.7	**114.9**	96.3	89.1
w/ DPC	**110.3**	109.0	98.5	**113.1**	**84.6**	**112.2**	**112.1**	**115.0**	**122.3**	**127.6**	**132.1**	97.1	106.2	**121.8**	**107.7**	**107.1**	101.6	**99.2**	86.8
w/o LRL	94.2	82.5	**102.1**	85.8	53.0	100.3	101.1	100.7	97.7	98.1	125.9	93.4	**106.6**	108.7	102.0	103.6	92.9	96.7	85.1
w/ GAP	98.9	94.6	85.0	98.1	33.6	108.0	103.5	109.2	111.2	115.6	99.5	91.8	95.8	98.8	103.3	97.8	112.7	98.5	64.1

Table B.4: Relative CD for corrupted variants of the validation set of the Cityscapes dataset for several network backbones of the DeepLabv3+ architecture and respective architectural ablations. Highest rCD per corruption is bold.

		Blur				Noise					Digital				Weather				
Deeplab-v3+ Backbone	Clean	Motion	Defocus	Frosted Glass	Gaussian	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
ResNet-50	69.6	38.7	**43.5**	31.1	45.5	43.2	40.7	44.2	50.9	59.8	63.5	50.3	63.8	58.2	31.3	47.0	56.9	39.8	67.2
w/o ASPP	57.6	28.8	28.8	21.9	31.5	31.7	29.2	32.3	38.4	45.9	49.8	36.0	51.1	45.0	23.1	37.6	42.0	26.7	56.3
w/o AC	68.9	39.3	41.6	29.0	43.6	43.6	42.0	44.1	50.8	59.3	62.7	48.9	62.9	56.9	31.2	46.4	55.9	38.3	65.9
w/ DPC	68.0	38.6	40.6	29.7	**42.5**	44.0	42.2	45.2	51.6	59.8	61.5	48.9	62.6	56.5	30.4	46.8	56.0	37.9	64.9
w/o LRL	69.0	40.0	41.2	30.1	43.0	43.5	41.9	44.3	50.8	59.7	62.4	48.5	62.5	57.2	30.4	46.6	55.8	39.1	65.5
w/ GAP	**71.5**	**41.6**	42.9	**33.0**	**45.9**	**45.6**	**45.3**	**46.4**	**53.7**	**63.4**	**65.7**	**52.1**	**66.1**	**59.5**	**33.6**	**50.0**	**60.3**	**43.8**	**67.6**
ResNet-101	70.3	45.8	45.6	33.2	46.6	49.4	48.3	50.1	55.4	61.3	64.5	50.6	65.3	59.7	31.4	50.4	57.6	41.2	67.6
w/o ASPP	60.5	36.4	34.2	25.1	36.3	36.4	34.1	37.1	43.0	50.5	53.6	39.2	54.1	49.8	24.5	41.9	45.5	29.6	59.8
w/o AC	70.2	**46.8**	45.8	33.5	46.3	46.0	45.2	46.6	52.9	60.5	64.4	50.5	64.5	59.6	32.3	51.0	57.9	40.4	66.8
w/ DPC	69.5	44.5	44.8	32.3	46.2	48.4	45.0	49.4	54.8	61.5	63.5	51.3	64.0	59.4	32.3	49.9	58.3	40.5	65.7
w/o LRL	69.6	44.2	44.8	33.5	45.8	47.4	45.2	48.5	53.8	61.3	64.1	50.7	64.6	58.4	32.2	50.6	57.5	40.4	65.9
w/ GAP	**72.5**	46.7	**46.3**	**36.5**	**47.6**	**50.5**	**48.5**	**51.3**	**56.6**	**64.3**	**66.7**	**53.6**	**66.0**	**61.1**	**36.4**	**52.6**	**61.7**	**44.7**	**68.4**
Xception-41	75.5	52.9	54.7	35.5	53.9	55.8	53.3	56.7	62.8	67.6	70.8	51.9	70.9	64.6	42.5	59.0	63.1	48.4	**73.0**
w/o ASPP	66.9	45.9	45.3	30.4	45.6	47.2	45.5	48.0	52.9	58.5	61.0	43.1	61.5	56.0	34.6	50.6	53.1	39.3	65.4
w/o AC	75.0	53.2	54.9	36.5	54.9	54.1	52.6	55.5	61.4	67.1	69.7	50.5	70.5	64.5	40.9	60.1	62.3	47.0	71.8
w/ DPC	75.3	51.6	54.8	**37.5**	54.3	**56.8**	55.1	58.1	63.0	67.8	70.0	50.8	70.7	65.5	40.6	58.1	61.9	47.7	72.0
w/o LRL	76.1	52.9	**56.7**	36.7	**55.8**	56.7	**56.6**	**58.3**	**63.9**	68.8	70.9	**53.8**	**71.9**	65.1	41.4	59.1	63.3	48.4	72.9
w/ GAP	**76.5**	**55.0**	55.2	36.3	55.0	55.3	55.7	56.6	62.7	**68.8**	**71.1**	52.8	71.5	**66.1**	**43.3**	**61.4**	**63.7**	**48.9**	72.3
Xception-65	76.5	53.5	58.3	37.7	57.2	56.6	54.7	57.4	62.5	69.3	71.8	55.9	72.1	66.7	40.2	58.5	64.0	47.5	**73.6**
w/o ASPP	70.6	47.5	47.8	29.1	48.6	45.4	44.2	45.6	51.8	62.4	64.7	48.1	64.6	58.3	35.7	52.6	56.4	39.4	68.7
w/o AC	76.4	**57.6**	57.3	38.4	56.9	56.5	54.5	57.0	62.2	69.6	71.4	55.0	72.3	66.3	42.5	60.4	63.6	46.4	73.6
w/ DPC	76.1	53.7	55.0	34.6	55.0	54.8	54.0	56.0	61.6	68.9	70.9	54.0	71.1	66.0	40.9	58.3	61.9	46.5	73.4
w/o LRL	76.2	55.2	55.1	36.6	55.6	56.5	55.4	56.8	61.8	68.9	71.1	56.1	71.1	64.3	40.3	58.4	64.2	46.1	73.0
w/ GAP	**77.5**	56.8	**59.8**	**41.8**	**59.1**	**57.9**	**57.6**	**57.6**	**62.6**	**71.0**	**73.1**	**57.4**	**73.0**	**67.3**	**42.8**	**61.1**	**65.3**	**49.7**	73.2
Xception-71	76.7	56.5	59.1	40.2	59.5	56.6	57.8	57.6	63.2	69.9	72.1	57.1	72.6	68.1	43.9	60.9	66.1	50.9	**73.6**
w/o ASPP	70.5	48.3	49.2	33.3	50.1	47.5	47.1	48.2	54.6	62.7	65.1	48.8	65.6	60.2	37.0	53.4	57.3	44.1	69.3
w/o AC	75.7	55.9	58.8	**41.8**	59.0	57.1	58.2	57.3	62.6	69.5	71.0	56.9	71.4	67.6	41.9	60.9	64.1	48.2	73.0
w/ DPC	76.8	53.5	54.6	35.8	55.4	55.5	55.6	54.7	60.5	69.0	71.4	54.0	71.0	66.3	42.5	58.3	63.3	49.7	73.3
w/o LRL	76.3	56.4	56.4	40.5	55.9	**59.9**	**59.3**	**60.2**	**64.6**	71.1	72.1	53.0	72.3	67.7	43.8	59.3	64.1	50.8	73.2
w/ GAP	**77.7**	**57.8**	**58.7**	38.3	**59.1**	58.8	55.2	58.4	63.7	**71.9**	**73.8**	**60.4**	**73.9**	**69.2**	**46.9**	**61.6**	**67.9**	**53.5**	73.5

Table B.5: Mean IoU for clean and corrupted variants of the validation set of PASCAL VOC 2012 for several network backbones of the DeepLabv3+ architecture and respective architectural ablations. Every mIoU is averaged over all available severity levels, except for corruptions of category noise where only the first three severity levels are considered. The standard deviation for image corruptions of category noise is 0.3 or less. Highest mIoU per corruption is bold.

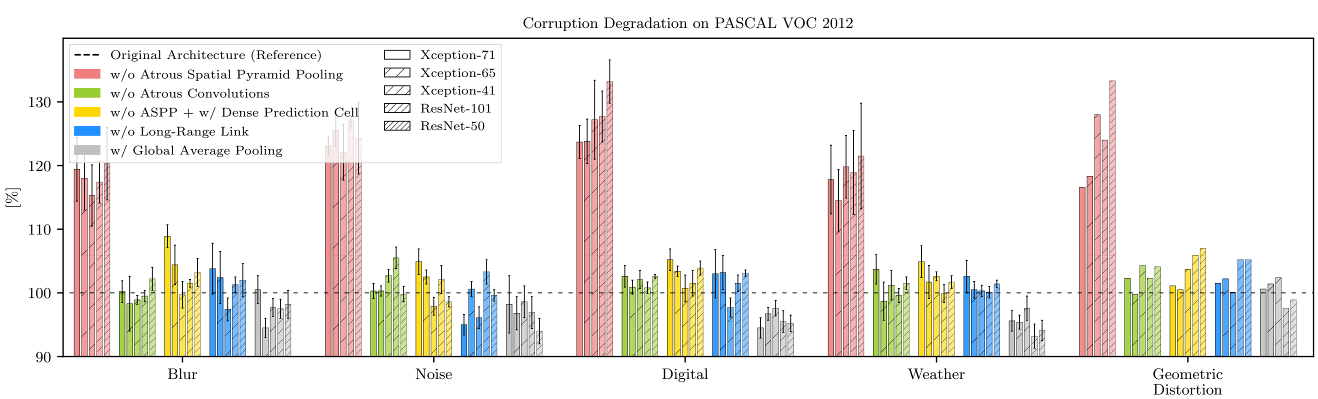


Figure B.10: CD evaluated on PASCAL VOC 2012 for the proposed ablated variants of the DeepLabv3+ architecture w.r.t. image corruptions, employing five different network backbones. Each bar except for geometric distortion is averaged within a corruption category (error bars indicate the standard deviation). Bars above 100 % represent a decrease in performance compared to the respective reference architecture. Each ablated architecture is re-trained on the original training dataset. Removing ASPP reduces the model performance significantly. AC and LRL decrease robustness against corruptions of category *digital* slightly. Xception-71 is vulnerable against many corruptions when DPC is used. GAP increases performance against many corruptions. Each backbone performs further best on clean data when GAP is used. Best viewed in color.

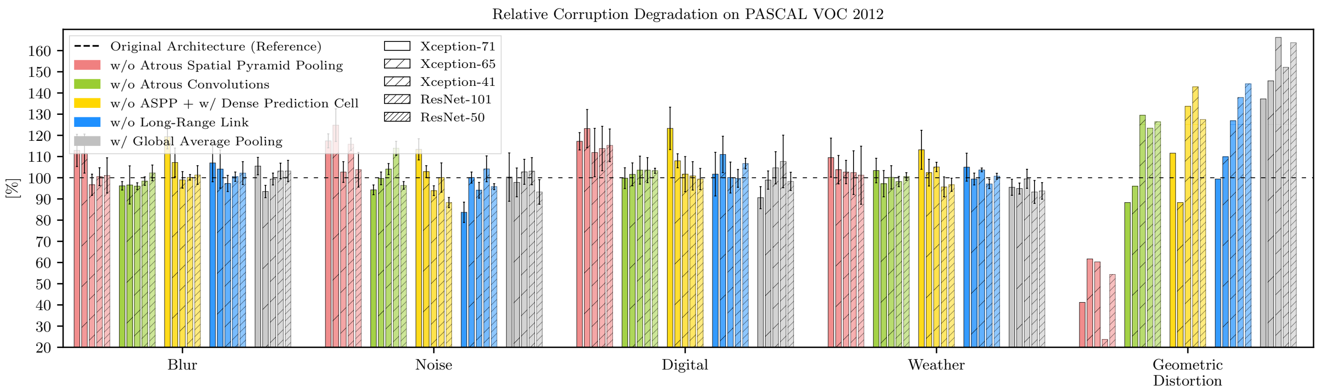


Figure B.11: Relative CD evaluated on PASCAL VOC 2012 for the proposed ablated variants of the DeepLabv3+ architecture w.r.t. image corruptions, employing five different network backbones. Each bar except for geometric distortion is averaged within a corruption category (error bars indicate the standard deviation). Bars above 100 % represent a relative decrease in performance compared to the respective reference architecture. Each ablated architecture is re-trained on the original training dataset. Removing ASPP decreases performance oftentimes significantly. The low rCD for geometric distortion indicates that the relative decrease of performance for this ablated variant is low. AC aids the robustness against geometric distortion for several backbones. The harming effect of DPC with respect to image corruptions is especially pronounced for Xception-71. The rCD of LRL is large against geometric distortion for ResNet-50. The rCD of GAP has, oftentimes, a contrary tendency as the CD. Best viewed in color.

	Blur				**Noise**					**Digital**				**Weather**				
Deeplab-v3+ Backbone	Motion	Defocus	Frosted Glass	Gaussian	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
ResNet-50	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**116.1**	**126.1**	**113.3**	**125.7**	**120.3**	**119.4**	**121.4**	**125.6**	**134.7**	**137.7**	**128.7**	**135.0**	**131.4**	**111.9**	**117.8**	**134.5**	**121.7**	**133.3**
w/o AC	99.1	103.4	103.0	103.5	99.3	97.9	100.2	100.4	101.3	102.3	102.6	102.4	103.1	100.0	101.1	102.2	102.6	104.1
w/ DPC	100.2	105.2	101.9	105.5	98.6	97.6	98.2	98.6	100.1	105.5	102.6	103.2	104.0	101.3	100.5	101.9	103.3	107.0
w/o LRL	98.0	104.1	101.5	104.6	99.4	98.0	99.9	100.3	100.3	103.0	103.6	103.4	102.4	101.2	100.9	102.4	101.2	105.2
w/ GAP	95.4	101.1	97.2	99.4	95.9	92.3	96.2	94.3	91.2	94.0	96.3	93.7	96.7	96.6	94.4	92.1	93.5	98.9
ResNet-101	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**117.2**	**120.9**	**112.2**	**119.4**	**125.7**	**127.4**	**126.1**	**127.8**	**127.9**	**130.9**	**123.0**	**132.3**	**124.5**	**110.1**	**117.1**	**128.5**	**119.8**	**124.0**
w/o AC	98.1	99.6	99.5	100.7	106.7	106.0	107.0	105.7	102.1	100.5	100.1	102.3	100.4	98.7	98.8	99.3	101.4	102.3
w/ DPC	102.3	101.4	101.4	100.8	102.0	106.2	101.5	101.4	99.5	103.0	98.6	103.6	100.9	98.7	101.2	98.3	101.3	105.9
w/o LRL	102.9	101.4	99.6	101.5	103.9	105.9	103.3	103.5	100.0	101.2	99.7	102.0	103.3	99.0	99.6	100.2	101.5	105.2
w/ GAP	98.3	98.6	95.0	98.2	97.8	99.5	97.7	97.2	92.2	93.8	93.9	97.8	96.6	92.8	95.6	90.4	94.0	97.6
Xception-41	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**114.8**	**120.8**	**107.9**	**117.9**	**119.3**	**116.6**	**120.1**	**126.5**	**127.9**	**133.7**	**118.3**	**132.3**	**124.4**	**113.6**	**120.6**	**127.2**	**117.6**	**128.0**
w/o AC	99.4	99.7	98.5	97.8	103.9	101.5	103.0	103.6	101.6	103.7	103.0	101.4	100.2	102.7	97.3	102.2	102.7	104.3
w/ DPC	102.9	99.8	97.0	99.1	97.8	96.1	96.8	99.4	99.4	102.6	102.3	100.5	97.4	103.2	102.3	103.2	101.5	103.7
w/o LRL	100.0	95.6	98.2	95.8	97.9	92.9	96.5	97.0	96.1	99.6	96.0	96.6	98.6	101.8	99.8	99.5	100.0	100.1
w/ GAP	95.5	99.0	98.8	97.5	101.2	94.9	100.3	100.3	96.4	99.0	98.1	97.8	95.7	98.6	94.3	98.4	99.0	102.4
Xception-65	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**112.8**	**125.2**	**113.9**	**120.2**	**125.9**	**123.1**	**127.5**	**128.7**	**122.3**	**125.3**	**117.7**	**126.7**	**125.3**	**107.4**	**114.0**	**121.1**	**115.4**	**118.3**
w/o AC	91.1	102.5	98.9	100.7	100.3	100.5	100.9	101.0	98.8	101.3	102.2	99.1	101.2	96.2	95.4	101.2	102.1	99.8
w/ DPC	99.5	108.0	104.9	105.3	104.1	101.5	103.2	102.6	101.2	103.4	104.4	103.5	102.3	98.8	100.3	105.8	101.9	100.5
w/o LRL	96.3	107.7	101.8	103.8	100.3	98.4	101.2	101.9	101.2	102.6	99.7	103.4	107.1	99.7	100.3	99.4	102.7	102.2
w/ GAP	92.8	96.4	93.4	95.6	97.0	93.6	99.4	99.8	94.3	95.3	96.6	96.7	98.2	95.6	93.6	96.5	95.7	101.4
Xception-71	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**118.9**	**124.3**	**111.5**	**123.0**	**121.0**	**125.2**	**122.0**	**123.4**	**123.8**	**125.0**	**119.2**	**125.6**	**124.9**	**112.3**	**119.1**	**126.0**	**113.9**	**116.6**
w/o AC	101.3	100.9	97.3	101.2	98.8	99.0	100.7	101.8	101.4	103.9	100.4	104.4	101.5	103.5	100.1	105.9	105.4	102.3
w/ DPC	107.0	111.1	107.4	110.1	102.5	105.1	106.7	107.5	102.9	102.4	107.1	105.8	105.5	102.5	106.7	108.1	102.5	101.1
w/o LRL	100.2	106.7	99.5	108.7	92.4	96.5	93.9	96.3	95.9	99.9	109.4	101.2	101.2	100.2	104.0	105.8	100.2	101.5
w/ GAP	97.1	101.1	103.2	100.8	94.8	106.1	98.2	98.7	93.2	93.9	92.3	95.2	96.5	94.7	98.3	94.7	94.7	100.6

Table B.6: Mean CD for corrupted variants of the validation set of PASCAL VOC 2012 for several network backbones of the DeepLabv3+ architecture and respective architectural ablations. Highest CD per corruption is bold.

	Blur				Noise					Digital				Weather				
Deeplab-v3+ Backbone	Motion	Defocus	Frosted Glass	Gaussian	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
ResNet-50	**100.0**	100.0	100.0	100.0	**100.0**	**100.0**	**100.0**	**100.0**	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0
w/o ASPP	93.1	110.4	92.6	108.2	98.3	98.3	99.8	103.0	**119.8**	**128.6**	**111.6**	**111.1**	**109.9**	90.0	88.7	**122.6**	**103.6**	54.4
w/o AC	95.9	104.7	**103.6**	105.0	95.8	93.2	97.7	97.2	98.2	102.1	103.2	102.8	105.2	98.2	99.5	102.0	102.8	126.4
w/ DPC	95.1	105.0	99.2	105.7	90.7	89.4	89.7	87.7	83.8	106.3	98.4	92.2	100.5	98.1	94.0	93.6	101.2	127.5
w/o LRL	93.9	106.4	100.9	107.7	96.3	93.7	97.1	97.4	94.7	107.4	106.0	110.1	103.1	**100.5**	99.1	103.0	100.2	144.3
w/ GAP	96.9	**109.5**	99.8	**106.4**	98.2	90.8	99.1	95.2	83.2	95.3	100.1	93.5	104.5	98.7	95.1	88.1	93.2	**163.7**
ResNet-101	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	**100.0**	100.0	100.0	100.0
w/o ASPP	98.0	**106.3**	95.4	102.3	115.3	**119.7**	115.8	**117.3**	111.0	**119.4**	**107.8**	127.7	100.5	92.6	93.3	**117.7**	**106.2**	23.7
w/o AC	95.4	98.5	98.7	101.1	**115.7**	113.4	**116.6**	116.1	**107.8**	100.9	99.5	113.7	100.1	97.4	96.4	96.8	102.4	123.4
w/ DPC	102.0	99.9	**100.4**	98.5	101.0	111.1	99.9	99.1	89.4	105.3	92.7	109.4	96.1	95.8	99.1	88.1	99.9	142.9
w/o LRL	103.7	100.3	97.4	100.6	106.3	110.9	105.0	106.1	92.7	96.1	95.9	100.7	106.3	96.5	95.6	95.3	100.8	137.9
w/ GAP	**105.1**	105.7	96.8	**105.1**	105.0	108.7	105.2	106.3	90.6	99.6	95.7	**127.9**	**107.4**	92.9	99.9	85.0	95.4	**152.1**
Xception-41	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	93.0	**104.0**	91.3	98.6	99.9	96.4	100.7	**110.4**	**106.0**	**127.2**	101.1	**118.5**	100.7	97.8	99.3	**111.8**	102.0	60.3
w/o AC	96.9	97.3	96.6	93.4	106.5	**101.1**	104.5	107.1	101.1	113.9	**104.2**	99.4	96.7	103.4	90.6	103.1	**103.6**	129.5
w/ DPC	**105.3**	98.9	94.8	97.2	94.2	91.0	91.7	97.0	95.5	112.9	103.9	99.9	89.9	**105.2**	**104.7**	108.4	102.2	133.7
w/o LRL	102.9	93.5	98.8	94.0	98.5	87.9	95.4	96.2	92.4	111.5	94.6	92.8	**101.3**	105.1	103.5	103.7	102.5	126.9
w/ GAP	95.0	102.8	**100.6**	99.3	**107.8**	93.7	**106.0**	108.7	97.9	115.4	100.4	107.8	95.3	100.6	92.0	103.5	101.9	**166.2**
Xception-65	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	**100.0**	100.0	100.0	100.0
w/o ASPP	**100.1**	**124.9**	106.9	**113.8**	**126.4**	**120.7**	**130.2**	**134.3**	**112.4**	**124.8**	**109.0**	**133.9**	**125.0**	95.8	99.3	**113.0**	**107.3**	61.7
w/o AC	81.7	105.2	98.1	101.3	100.4	100.6	101.7	102.0	93.8	106.5	104.3	92.5	103.3	93.5	89.0	103.0	103.5	96.0
w/ DPC	96.9	115.7	**106.7**	109.3	106.5	101.1	104.7	103.5	98.4	110.5	107.2	111.3	102.9	96.7	98.1	112.9	101.8	88.3
w/o LRL	91.2	116.1	102.2	106.9	99.3	95.5	101.3	103.1	101.1	109.5	98.0	115.1	121.4	98.7	99.1	96.0	103.8	109.9
w/ GAP	89.7	97.1	91.9	95.4	98.5	91.1	103.8	106.6	89.3	92.7	97.5	101.6	103.9	95.5	90.6	97.8	95.6	**145.7**
Xception-71	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	110.1	121.4	101.9	118.3	**114.5**	**123.5**	**116.5**	118.0	**114.6**	**117.1**	110.6	120.5	120.6	102.2	108.2	124.9	102.6	41.2
w/o AC	98.1	96.6	92.9	97.3	92.7	92.7	96.6	97.7	91.9	102.8	96.0	106.3	94.3	103.1	94.3	109.8	**106.6**	88.3
w/ DPC	**115.3**	**126.1**	**112.2**	**124.2**	105.6	111.6	115.1	**120.9**	113.5	115.8	115.9	**140.2**	**121.2**	**104.4**	**116.9**	**126.6**	104.9	**111.6**
w/o LRL	98.4	113.3	98.1	118.1	81.6	90.0	84.3	86.8	75.8	90.7	**118.6**	97.8	99.8	99.0	107.4	114.8	98.7	99.4
w/ GAP	98.5	108.2	107.9	107.5	93.7	118.8	101.0	103.7	84.2	84.3	88.1	91.6	98.3	93.9	102.0	92.3	93.7	137.2

Table B.7: Relative CD for corrupted variants of the validation set of PASCAL VOC 2012 for several network backbones of the DeepLabv3+ architecture and respective architectural ablations. Highest rCD per corruption is bold.

		Blur				Noise					Digital				Weather				
Deeplab-v3+ Backbone	Clean	Motion	Defocus	Frosted Glass	Gaussian	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
MobileNet-V2	33.1	16.1	16.6	14.9	16.5	12.1	11.5	12.4	17.0	24.7	27.2	14.8	26.5	25.1	7.8	18.5	20.1	10.7	28.3
w/o ASPP	27.3	12.2	11.3	10.5	11.6	9.6	9.6	9.9	13.3	19.0	22.1	10.8	20.8	19.5	5.7	15.6	14.5	7.8	22.9
w/o AC	32.1	15.2	15.9	14.2	15.7	11.2	11.5	11.4	15.6	23.0	27.1	13.7	25.1	25.1	7.6	18.8	19.2	10.7	27.9
w/ DPC	**34.7**	**17.3**	**18.9**	**15.6**	**18.0**	**13.9**	**13.7**	**13.8**	**18.4**	**25.8**	**28.9**	**15.8**	**27.8**	26.2	8.6	20.8	**21.5**	11.6	**29.4**
w/o LRL	32.2	15.7	16.5	14.3	16.3	12.9	11.6	13.2	17.3	24.4	26.7	14.2	25.1	24.6	7.5	18.9	19.7	10.7	27.6
w/ GAP	33.9	17.2	17.9	15.1	17.2	12.5	12.8	12.8	16.8	25.3	28.7	15.0	27.4	**26.6**	**8.7**	**20.9**	20.8	**11.7**	29.1
ResNet-50	37.4	18.0	19.7	16.9	19.2	14.1	12.8	14.4	19.4	28.5	31.1	18.0	30.1	29.5	8.8	21.5	23.9	13.6	32.9
w/o ASPP	29.7	13.5	13.8	11.6	13.5	11.1	10.1	11.6	15.5	21.6	24.8	13.3	23.4	22.7	6.7	17.0	17.6	9.9	25.5
w/o AC	36.5	18.2	19.3	16.6	18.7	13.7	11.8	13.8	18.8	27.5	30.2	17.3	29.1	28.7	7.9	20.4	23.3	12.8	31.3
w/ DPC	37.9	18.9	20.3	**17.5**	19.8	13.4	12.2	13.7	19.2	29.0	31.6	18.9	30.3	30.1	8.6	20.9	24.5	13.6	33.0
w/o LRL	36.6	18.3	19.8	16.1	18.8	13.6	12.2	13.8	18.9	27.4	31.0	19.1	30.1	29.3	8.1	21.2	24.3	13.4	32.0
w/ GAP	**38.2**	**19.3**	**21.1**	17.2	**19.9**	**15.5**	**12.8**	**15.8**	**21.3**	**30.4**	**32.9**	**19.5**	**31.7**	**30.8**	**9.9**	**23.2**	**25.8**	**14.9**	**33.0**
ResNet-101	38.1	19.1	20.6	17.3	19.8	15.4	14.6	15.7	20.7	28.8	31.6	19.7	31.2	31.4	10.2	22.9	25.6	14.0	32.8
w/o ASPP	30.7	14.3	14.1	12.8	14.2	13.3	11.8	13.7	17.7	23.4	25.9	14.4	24.7	24.1	7.3	18.5	18.8	10.7	26.2
w/o AC	37.3	18.3	19.9	16.9	19.0	14.4	14.4	14.7	19.4	27.5	31.4	18.1	30.1	30.5	9.4	22.9	24.6	13.6	32.2
w/ DPC	37.6	19.6	21.0	17.7	20.0	15.9	**15.1**	16.4	21.6	28.7	32.1	19.5	31.5	31.2	9.7	23.3	25.4	14.0	32.6
w/o LRL	37.5	18.9	20.5	17.7	19.9	16.5	14.6	**16.8**	**21.7**	29.0	31.6	19.8	30.7	30.1	9.8	22.2	25.9	14.0	32.2
w/ GAP	**39.3**	**20.2**	**21.7**	**17.9**	**20.6**	**15.9**	14.2	16.1	21.4	**29.9**	**33.2**	**20.4**	**32.8**	**32.8**	**10.8**	**23.3**	**27.0**	**15.6**	**34.2**
Xception-41	39.7	22.1	22.7	17.4	**20.8**	20.8	18.1	20.5	24.8	**33.7**	34.2	20.9	32.5	32.6	13.0	25.0	28.4	17.0	**34.4**
w/o ASPP	35.4	19.4	20.0	15.3	18.4	18.2	16.3	17.9	21.7	29.2	30.3	17.7	28.5	28.3	11.2	22.7	23.5	14.4	31.3
w/o AC	38.4	21.8	22.2	**17.7**	20.6	21.8	18.3	21.1	25.0	32.9	33.4	20.0	31.7	32.0	12.2	24.8	26.4	15.9	33.0
w/ DPC	**40.2**	21.9	22.5	17.4	20.8	20.2	17.5	19.6	23.9	33.3	**34.8**	20.3	**32.6**	**32.9**	13.8	25.6	26.9	17.4	34.5
w/o LRL	39.1	21.4	22.6	17.2	20.6	20.8	17.6	20.5	25.0	32.6	34.1	21.1	32.1	32.2	13.8	25.5	27.1	16.9	34.2
w/ GAP	39.0	**22.7**	**22.9**	17.5	21.0	**21.9**	**18.5**	**21.6**	**25.4**	33.1	34.1	**21.6**	32.3	32.6	**14.3**	**25.8**	**28.9**	**17.7**	34.1
Xception-65	41.4	23.4	25.2	18.9	22.7	23.2	19.8	22.9	27.1	35.4	36.1	23.5	34.8	**34.2**	14.8	27.7	30.0	18.4	35.6
w/o ASPP	40.2	21.4	23.3	18.1	21.6	20.4	16.7	20.1	24.7	33.0	34.1	21.4	32.6	31.4	12.1	25.3	27.4	15.6	35.1
w/o AC	40.0	22.7	24.4	18.6	22.1	23.6	20.9	22.8	26.4	34.6	35.1	23.4	33.9	33.3	13.9	27.2	29.1	17.8	35.0
w/ DPC	40.9	23.7	24.9	18.4	22.8	23.0	18.8	22.9	27.0	35.6	35.7	23.0	34.1	33.9	14.6	28.0	29.5	17.9	35.7
w/o LRL	41.0	23.2	25.0	18.6	22.7	24.2	19.8	23.9	27.6	35.6	35.7	23.5	34.1	33.8	14.9	27.5	29.3	19.0	**36.1**
w/ GAP	**41.7**	**23.9**	**25.6**	**19.2**	**23.4**	**24.7**	**20.8**	**24.2**	**28.1**	**36.2**	**36.1**	**23.8**	**35.0**	34.1	**15.4**	**28.2**	**30.5**	**20.1**	36.0
Xception-71	42.4	**24.4**	26.4	19.5	23.9	**24.0**	**20.3**	**23.3**	**27.5**	**36.8**	**37.2**	**25.3**	**35.7**	**34.7**	16.1	**29.4**	31.3	19.8	**37.1**
w/o ASPP	40.6	21.9	24.2	17.5	21.9	20.8	16.6	20.0	24.2	34.0	34.8	22.5	33.1	32.4	12.9	26.3	28.9	16.5	35.2
w/o AC	41.8	24.3	25.4	**19.6**	23.6	24.0	19.9	22.9	26.2	35.7	36.1	23.2	34.8	33.7	15.7	28.3	29.9	19.7	35.8
w/ DPC	**42.5**	23.3	25.9	18.4	23.1	23.4	18.9	22.4	26.5	36.3	36.5	24.1	34.9	34.2	15.8	28.3	30.6	18.6	36.3
w/o LRL	42.2	22.9	25.9	18.7	23.5	21.7	18.7	21.1	25.4	35.5	36.3	24.3	34.5	34.0	15.1	28.6	30.6	19.7	36.4
w/ GAP	42.0	24.0	**26.6**	19.1	**24.0**	23.6	19.8	22.8	26.7	35.9	37.0	25.0	35.2	34.6	**16.7**	29.3	**31.6**	**20.9**	36.3

Table B.8: Mean IoU for clean and corrupted variants of the validation set of ADE20K for several network backbones of the DeepLabv3+ architecture and respective architectural ablations. Every mIoU is averaged over all available severity levels, except for corruptions of category noise where only the first three severity levels are considered. The standard deviation for image corruptions of category noise is 0.2 or less. Highest mIoU per corruption is bold.

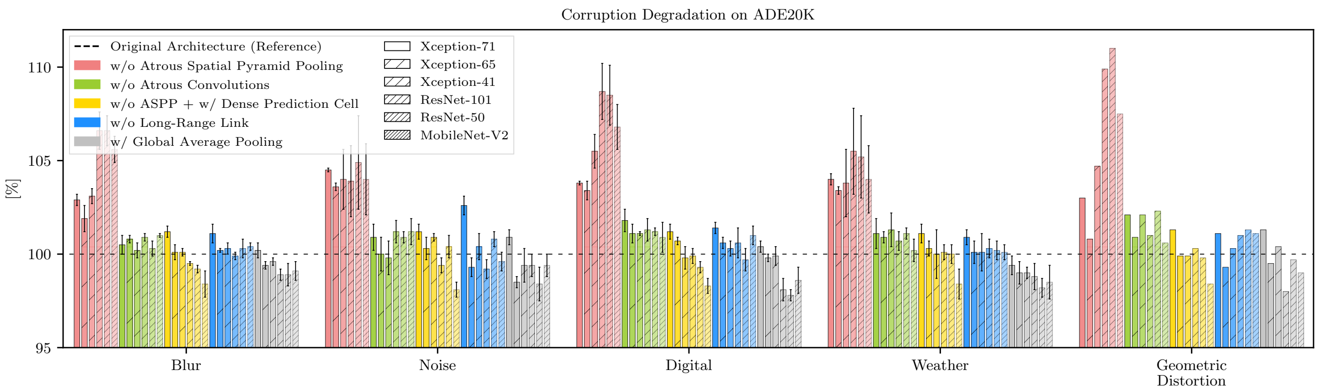


Figure B.12: CD evaluated on ADE20K for the proposed ablated variants of the DeepLabv3+ architecture w.r.t. image corruptions, employing six different network backbones. Each bar except for geometric distortion is averaged within a corruption category (error bars indicate the standard deviation). Bars above 100 % represent a relative decrease in performance compared to the respective reference architecture. Each ablated architecture is re-trained on the original training dataset. Removing ASPP decreases performance oftentimes. AC increase performance slightly against most corruptions. DPC and LRL hamper the performance for Xception-71 w.r.t. several image corruptions. GAP increases the robustness for most backbones against many image corruptions. Best viewed in color.

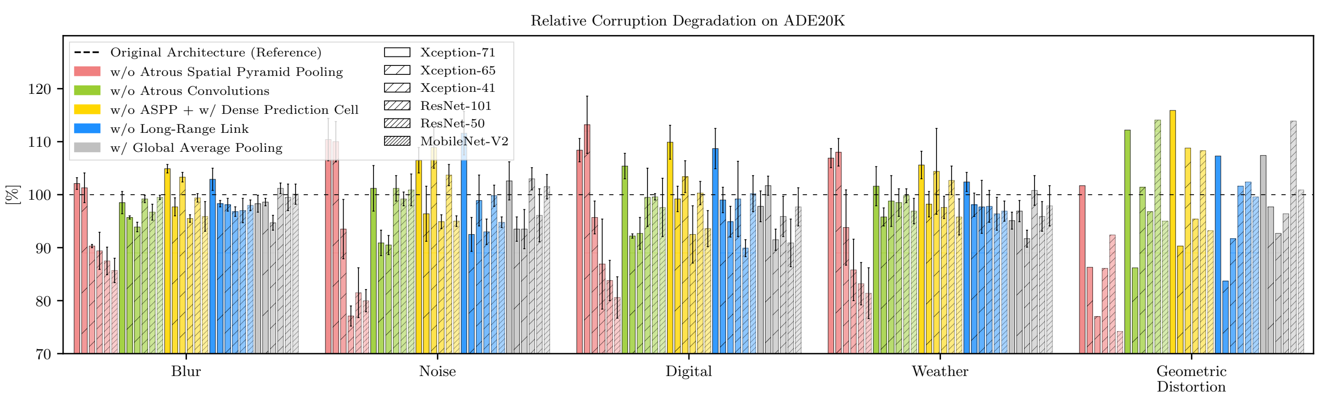


Figure B.13: Relative CD evaluated on ADE20K for the proposed ablated variants of the DeepLabv3+ architecture w.r.t. image corruptions, employing six different network backbones. Bars above 100 % represent a relative decrease in performance compared to the respective reference architecture. Each bar except for geometric distortion is averaged within a corruption category (error bars indicate the standard deviation). Each ablated architecture is re-trained on the original training dataset. Removing ASPP decreases performance oftentimes significantly. The low rCD for geometric distortion indicates that the relative decrease of performance for this ablated variant is low (except for Xception-71). The rCD of DPC and LRL are oftentimes highest for Xception-71. GAP increases the robustness for most backbones against many image corruptions. Best viewed in color.

	Blur				**Noise**					**Digital**				**Weather**				
Deeplab-v3+ Backbone	Motion	Defocus	Frosted Glass	Gaussian	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
MobileNet-V2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**104.7**	**106.4**	**105.2**	**105.9**	**102.9**	**102.1**	**102.8**	**104.5**	**107.5**	**107.0**	**104.8**	**107.8**	**107.5**	**102.3**	**103.6**	**107.1**	**103.3**	**107.5**
w/o AC	101.1	100.9	100.8	101.0	101.1	100.0	101.1	101.7	102.1	100.1	101.3	102.0	100.0	100.2	99.6	101.1	100.0	100.6
w/ DPC	98.7	97.3	99.2	98.3	98.0	97.5	98.4	98.3	98.5	97.7	98.8	98.3	98.5	99.2	97.2	98.2	99.0	98.4
w/o LRL	100.6	100.1	100.8	100.3	99.1	99.8	99.1	99.6	100.3	100.8	100.8	102.0	100.7	100.4	99.5	100.6	100.0	101.1
w/ GAP	98.8	98.5	99.8	99.3	99.6	98.5	99.5	100.2	99.1	98.0	99.8	98.8	98.0	99.0	97.0	99.2	98.9	99.0
ResNet-50	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**105.4**	**107.4**	**106.4**	**107.1**	**103.5**	**103.1**	**103.3**	**104.8**	**109.7**	**109.2**	**105.8**	**109.5**	**109.6**	**102.2**	**105.8**	**108.3**	**104.3**	**111.0**
w/o AC	99.7	100.6	100.5	100.6	100.5	101.1	100.8	100.6	101.4	101.3	100.9	101.4	101.2	101.0	101.5	100.8	100.9	102.3
w/ DPC	98.9	99.3	99.3	99.4	100.8	100.7	100.9	100.2	99.3	99.3	98.9	99.7	99.2	100.1	100.8	99.2	100.0	99.8
w/o LRL	99.6	99.9	101.0	100.6	100.6	100.7	100.7	100.5	101.6	100.1	98.7	99.9	100.3	100.7	100.4	99.5	100.3	101.3
w/ GAP	98.4	98.3	99.7	99.2	98.4	100.0	98.4	97.7	97.4	97.4	98.1	97.7	98.1	98.7	97.9	97.5	98.5	99.7
ResNet-101	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**105.9**	**108.1**	**105.5**	**107.0**	**102.5**	**103.2**	**102.3**	**103.8**	**107.5**	**108.4**	**106.5**	**109.4**	**110.6**	**103.2**	**105.7**	**109.2**	**103.9**	**109.9**
w/o AC	101.0	100.8	100.5	101.0	101.2	100.2	101.2	101.6	101.8	100.3	101.9	101.5	101.3	100.9	100.0	101.4	100.5	101.0
w/ DPC	99.3	99.4	99.6	99.7	99.4	99.3	99.1	98.9	100.1	99.3	100.3	99.6	100.4	100.6	99.5	100.3	100.1	100.3
w/o LRL	100.2	100.1	99.6	99.9	98.7	100.0	98.7	98.8	99.6	100.0	99.9	100.7	101.9	100.5	100.9	99.6	100.0	101.0
w/ GAP	98.7	98.6	99.3	98.9	99.5	100.4	99.5	99.2	98.3	97.6	99.1	97.6	97.9	99.4	99.5	98.1	98.1	98.0
Xception-41	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**103.4**	**103.4**	**102.5**	**103.0**	**103.4**	**102.3**	**103.3**	**104.1**	**106.9**	**105.9**	**104.0**	**105.9**	**106.4**	**102.1**	**103.1**	**106.9**	**103.1**	**104.7**
w/o AC	100.3	100.6	99.6	100.3	98.7	99.8	99.2	99.8	101.3	101.2	101.1	101.1	100.9	100.9	100.3	102.8	101.2	102.1
w/ DPC	100.3	100.2	99.9	100.0	100.8	100.8	101.2	101.2	100.7	99.1	100.8	99.8	99.6	99.1	99.3	102.2	99.4	99.9
w/o LRL	100.8	100.1	100.2	100.2	100.0	100.7	100.0	99.7	101.7	100.1	99.8	100.6	100.6	99.1	99.4	101.8	100.0	100.3
w/ GAP	99.2	99.6	99.9	99.7	98.6	99.5	98.5	99.2	101.0	100.1	99.1	100.3	100.0	98.5	99.0	99.4	99.2	100.4
Xception-65	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**102.7**	**102.6**	**101.0**	**101.5**	**103.5**	**104.0**	**103.7**	**103.2**	**103.6**	**103.2**	**102.8**	**103.4**	**104.2**	**103.2**	**103.3**	**103.8**	**103.5**	100.8
w/o AC	100.9	101.0	100.4	100.8	99.4	98.7	100.1	100.9	101.1	101.6	100.2	101.3	101.2	101.0	100.7	101.3	100.7	**100.9**
w/ DPC	99.6	100.3	100.7	99.9	100.2	101.3	100.1	100.2	99.6	100.6	100.6	101.1	100.4	100.2	99.6	100.7	100.6	99.9
w/o LRL	100.3	100.2	100.4	100.0	98.6	100.1	98.8	99.2	99.6	100.5	100.1	101.1	100.6	99.9	100.2	101.0	99.3	99.3
w/ GAP	99.4	99.4	99.7	99.2	98.0	98.8	98.3	98.6	98.7	99.9	99.6	99.7	100.1	99.3	99.3	99.3	97.9	99.5
Xception-71	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	**103.3**	**103.1**	**102.6**	**102.6**	**104.3**	**104.6**	**104.3**	**104.6**	**104.5**	**103.8**	**103.8**	**103.9**	**103.6**	**103.8**	**104.4**	**103.6**	**104.1**	**103.0**
w/o AC	100.2	101.4	100.0	100.4	100.0	100.5	100.6	101.8	101.8	101.7	102.8	101.3	101.5	100.5	101.5	102.1	100.1	102.1
w/ DPC	101.4	100.8	101.4	101.0	100.7	101.8	101.2	101.4	100.9	101.1	101.7	101.3	100.7	100.4	101.5	101.1	101.5	101.3
w/o LRL	101.9	100.7	101.0	100.6	103.0	102.0	103.0	102.9	102.1	101.3	101.5	101.9	101.1	101.3	101.0	101.0	100.2	101.1
w/ GAP	100.5	99.7	100.5	99.9	100.5	100.6	100.8	101.1	101.5	100.2	100.4	100.8	100.1	99.3	100.1	99.6	98.6	101.3

Table B.9: CD for corrupted variants of the validation set of ADE20K for several network backbones of the DeepLabv3+ architecture and respective architectural ablations. Highest CD per corruption is bold.

	Blur				**Noise**					**Digital**				**Weather**				
Deeplab-v3+ Backbone	Motion	Defocus	Frosted Glass	Gaussian	Gaussian	Impulse	Shot	Speckle	Intensity	Brightness	Contrast	Saturate	JPEG	Snow	Spatter	Fog	Frost	Geometric Distortion
MobileNet-V2	100.0	**100.0**	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0	**100.0**	100.0	100.0	**100.0**	**100.0**	**100.0**	100.0	**100.0**	100.0
w/o ASPP	82.3	88.6	85.4	86.6	79.5	77.4	78.9	80.2	83.8	74.0	83.8	81.9	82.9	81.0	74.5	88.0	82.0	74.2
w/o AC	**100.1**	99.4	98.9	99.6	100.0	96.5	100.2	102.4	**105.5**	91.8	100.9	105.0	92.8	97.7	93.3	100.0	96.7	95.0
w/ DPC	97.2	91.7	99.2	95.5	95.1	93.4	96.4	95.5	94.9	88.5	97.8	93.2	95.0	99.2	90.7	94.8	98.5	93.2
w/o LRL	98.5	96.7	99.4	97.5	93.6	96.3	93.8	94.5	95.7	97.7	99.3	**106.0**	97.7	98.6	93.8	98.1	97.1	99.6
w/ GAP	99.0	97.8	**102.8**	**100.9**	**101.7**	97.9	**101.7**	**105.0**	101.4	94.0	**102.8**	99.3	94.6	99.8	91.4	**101.0**	99.3	**100.9**
ResNet-50	**100.0**	**100.0**	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0	**100.0**	100.0	100.0	**100.0**	100.0	100.0
w/o ASPP	83.1	89.8	88.1	89.0	79.6	79.7	78.6	78.6	90.9	77.6	84.5	85.4	87.9	80.2	79.9	89.8	83.1	92.4
w/o AC	94.1	97.4	97.4	97.8	98.0	100.1	98.9	97.8	101.1	100.0	99.0	100.4	99.0	99.9	101.6	98.0	99.6	114.1
w/ DPC	98.0	99.6	99.8	100.0	**105.0**	**104.5**	**105.5**	**103.9**	99.8	**100.5**	98.1	**103.8**	98.9	**102.3**	**107.0**	99.4	**102.1**	108.3
w/o LRL	94.3	95.1	100.0	98.4	98.8	99.4	99.2	98.0	**103.8**	88.5	90.4	88.3	92.4	99.6	97.1	91.4	97.6	102.4
w/ GAP	97.3	97.1	**102.9**	**100.8**	97.6	103.2	97.6	94.1	88.1	84.4	96.4	89.5	93.5	98.9	94.7	91.8	98.0	**113.9**
ResNet-101	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0
w/o ASPP	86.3	94.7	86.3	90.3	76.9	80.3	75.8	74.7	77.9	74.6	88.3	86.3	98.5	83.9	80.6	95.7	83.2	86.1
w/o AC	99.9	99.1	98.2	99.8	100.8	97.1	100.8	102.7	**104.5**	90.1	**103.8**	**102.8**	**101.3**	99.9	94.4	**101.4**	98.2	96.8
w/ DPC	94.8	94.8	96.0	96.3	95.7	95.6	94.7	92.4	96.1	85.5	98.6	89.1	96.8	100.1	94.2	97.8	98.3	95.4
w/o LRL	98.0	97.1	95.6	96.4	92.6	97.6	92.6	91.3	90.8	91.2	96.5	98.6	110.6	99.6	101.1	93.0	97.6	**101.6**
w/ GAP	**100.5**	**100.0**	**102.6**	**101.6**	**103.1**	**106.2**	**103.2**	**102.9**	99.6	92.8	102.3	92.9	95.6	**102.3**	**104.8**	98.2	98.1	96.4
Xception-41	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0	**100.0**	**100.0**	100.0	100.0	100.0
w/o ASPP	90.8	90.3	90.2	89.8	91.4	88.7	91.1	91.9	104.6	92.7	93.8	95.5	100.7	90.7	86.6	105.6	92.3	77.0
w/o AC	94.0	95.1	92.5	94.1	87.8	92.9	89.7	90.1	92.1	90.7	97.7	92.2	90.2	98.0	92.5	106.1	98.7	101.4
w/ DPC	**104.3**	**104.2**	**102.2**	**102.6**	**106.1**	**105.4**	**107.6**	**109.5**	**116.2**	98.5	**106.0**	**105.5**	**103.4**	99.0	100.1	**118.5**	**100.2**	**108.8**
w/o LRL	100.0	96.9	98.0	97.4	96.5	99.5	96.5	94.2	107.8	89.9	95.6	97.0	97.1	94.7	92.8	105.9	97.3	91.7
w/ GAP	92.8	94.4	96.6	95.1	90.7	95.1	90.4	91.3	100.1	88.6	92.6	93.8	90.9	92.5	90.3	90.2	94.0	92.7
Xception-65	100.0	100.0	**100.0**	**100.0**	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**100.0**
w/o ASPP	**104.3**	**103.8**	97.9	99.3	**107.8**	**108.8**	**108.5**	**107.5**	**117.5**	**113.6**	**104.9**	**114.0**	**120.2**	**105.3**	**108.0**	**112.2**	**106.6**	86.3
w/o AC	96.0	95.8	95.1	95.8	89.8	88.7	92.8	94.7	88.5	92.3	92.8	91.9	91.8	98.0	93.3	95.6	96.3	86.2
w/ DPC	95.4	98.3	100.0	96.9	97.9	102.6	97.6	97.1	86.7	97.4	99.8	102.8	96.7	98.7	94.1	100.0	99.8	90.3
w/o LRL	98.6	97.9	99.2	97.5	91.4	98.0	92.2	92.6	88.2	97.0	97.6	103.0	98.5	97.9	97.5	101.6	95.4	83.7
w/ GAP	98.6	98.6	99.7	97.7	92.7	96.6	94.2	94.4	89.7	102.8	99.6	100.4	104.0	98.7	97.9	97.6	93.5	97.7
Xception-71	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
w/o ASPP	103.6	102.7	101.1	100.8	107.6	**108.3**	107.8	109.9	118.3	**110.7**	106.0	110.5	**106.5**	**105.3**	**109.8**	105.8	**106.6**	101.7
w/o AC	97.1	102.1	97.0	98.0	96.2	98.7	98.9	104.1	108.0	107.2	**108.3**	102.6	103.7	99.2	103.0	107.0	97.3	112.2
w/ DPC	106.0	**103.9**	**105.1**	**104.4**	103.2	106.7	105.1	107.0	110.6	113.7	107.4	112.5	106.1	101.4	108.2	**107.2**	105.5	**115.9**
w/o LRL	**106.5**	101.7	102.3	101.0	**110.7**	105.9	**110.4**	**112.4**	**118.5**	110.3	104.8	**114.1**	105.5	103.0	103.5	103.7	99.4	107.3
w/ GAP	99.7	96.1	100.0	97.3	100.0	100.3	100.9	102.3	109.8	94.5	99.5	101.6	95.4	96.1	97.3	93.8	93.2	107.4

Table B.10: Relative CD for corrupted variants of the validation set of ADE20K for several network backbones of the DeepLabv3+ architecture and respective architectural ablations. Highest rCD per corruption is bold.