Supplemental Material CVPR2022 Paper
Unifying Panoptic Segmentation for Autonomous Driving

1. Unified Label Policy

Table 1 lists all WD2 labels including mappings to MVD, IDD, Cityscapes, and WD2_{eval} as well as each label's supercategory and visualization color.

WD2	color	MVD	IDD	Cityscapes	WD2_{eval}
person		**person**	**person**	**person**	**person**
motorcyclist		**motorcyclist**	**rider**	**rider**	**rider**
bicyclist		**bicyclist**	**rider**	**rider**	**rider**
otherrider		**otherrider**	**rider**	**rider**	**rider**
egovehicle		egovehicle	**egovehicle**	*egovehicle*	**egovehicle**
dashcammount		carmount	**egovehicle**	*egovehicle*	**egovehicle**
car		**car**	**car**	**car**	**car**
truck		**truck**	**truck**	**truck**	**truck**
bus		**bus**	**bus**	**bus**	**bus**
motorcycle		**motorcycle**	**motorcycle**	**motorcycle**	**motorcycle**
bicycle		**bicycle**	**bicycle**	**bicycle**	**bicycle**
pickup		**truck**	**truck**	**truck**	**pickup**
van		**car**	**car**	**car**	**van**
autorickshaw		**othervehicle**	**autorickshaw**	**motorcycle**	**motorcycle**
caravan		**caravan**	**caravan**	***caravan***	*unlabeled*
trailer		**trailer**	**trailer**	***trailer***	*unlabeled*
onrails		onrails	**train**	**onrails**	*unlabeled*
othervehicle		**othervehicle**	**vehiclef.**	*dynamic*	*unlabeled*
wheeledslow		**wheeledslow**	**vehiclef.**	*dynamic*	*unlabeled*
boat		**boat**	*unlabeled*	*dynamic*	*unlabeled*
road		road	road	road	road
sidewalk		sidewalk	sidewalk	sidewalk	sidewalk
roadmarking		markinggeneral	road	road	roadmarking
curb		curb	curb	sidewalk	sidewalk
tramtrack		railtrack	road	road	road
bikelane		bikelane	road	road	road
bikelanesidewalk		bikelane	sidewalk	sidewalk	sidewalk
pedestrianarea		pedestrianarea	road	road	road
crosswalkplain		**crosswalkplain**	road	road	road
crosswalkzebra		**crosswalkzebra**	road	road	road
curbterrain		curb	curb	terrain	terrain
servicelane		servicelane	road	road	road
curbcut		curbcut	curb	sidewalk	sidewalk
ground		*unlabeled*	*unlabeled*	*ground*	*unlabeled*
parking		parking	parking	*parking*	*unlabeled*
railtrack		railtrack	railtrack	*railtrack*	*unlabeled*
water		water	nondrivablef.	*ground*	*unlabeled*
sand		sand	drivablef.	*ground*	*unlabeled*
snow		snow	*unlabeled*	*ground*	*unlabeled*
polegroup		**pole**	*polegroup*	*polegroup*	pole
pole		**pole**	pole	*pole*	pole
utilitypole		**utilitypole**	pole	*pole*	pole
trafficsignframe		**trafficsignframe**	pole	*pole*	pole
trafficlight		**trafficlight**	trafficlight	trafficlight	trafficlight
billboard		**billboard**	billboard	*billboard*	billboard
streetlight		**streetlight**	obsstrbarf.	*streetlight*	streetlight
manhole		**manhole**	road	road	road
trafficsign		**trafficsignfront**	trafficsign	trafficsign	trafficsign
trafficsignback		**trafficsignback**	obsstrbarf.	*static*	*unlabeled*
trafficsignany		*unlabeld*	obsstrbarf.	*static*	*unlabeled*
otherbarrier		otherbarrier	wall	wall	wall
catchbasin		**catchbasin**	road	road	road
manholesidewalk		**manhole**	sidewalk	sidewalk	sidewalk
junctionbox		**junctionbox**	obsstrbarf.	*static*	*unlabeled*
mailbox		**mailbox**	obsstrbarf.	*static*	*unlabeled*
phonebooth		**phonebooth**	obsstrbarf.	*static*	*unlabeled*
bikerack		**bikerack**	obsstrbarf.	*static*	*unlabeled*
pothole		pothole	road	road	road
trashcan		**trashcan**	obsstrbarf.	*static*	*unlabeled*
bench		**bench**	obsstrbarf.	*static*	*unlabeled*
banner		**banner**	obsstrbarf.	*dynamic*	*unlabeled*
firehydrant		**firehydrant**	obsstrbarf.	*static*	*unlabeled*
cctvcamera		**cctvcamera**	obsstrbarf.	*static*	*unlabeled*
building		building	building	building	building
wall		wall	wall	wall	wall
fence		fence	fence	fence	fence
guardrail		guardrail	guardrail	*guardrail*	guardrail
bridge		bridge	bridge	*bridge*	*unlabeled*
tunnel		tunnel	tunnel	*tunnel*	*unlabeled*
vegetation		vegetation	vegetation	vegetation	vegetation
terrain		terrain	nondrivablef.	terrain	terrain
groundanimal		**groundanimal**	**animal**	*dynamic*	*unlabeled*
bird		**bird**	**animal**	*dynamic*	*unlabeled*
mountain		mountain	f.background	*static*	*unlabeled*
sky		sky	sky	sky	sky
dynamic		*unlabeled*	*unlabeled*	*dynamic*	*unlabeled*
overlay		*unlabeled*	*rect.border*	*rect.border*	*unlabeled*
outofroi		*unlabeled*	*outofroi*	*outofroi*	*unlabeled*
static		*unlabeled*	*unlabeled*	*static*	*unlabeled*
unlabeled		*unlabeled*	*unlabeled*	*unlabeled*	*unlabeled*

Table 1. Wilddash2 label policy and mapping to MVD, IDD, CS, and WD2_{eval}. Bold labels have instance annotations, italic labels are not evaluated at their respective benchmark. Negative test cases do evaluate areas labeled as *unlabeled* in WD2_{eval} (see paper's Section 4.2 on Negative Testing); Supercategories: human; vehicle; flat; object; construction; nature; sky; void

2. Category definitions

The WD2 label policy unifies MVD, IDD, and Cityscapes category labels (in addition to the new vehicle labels *pickup* and *van*. The definition for most labels can be found in existing label definitions. Others need clarification or clear rules for differentiation in borderline cases. This leads to the following category definitions:

- The categories *person*, *egovehicle*, *car*, *truck*, *bus*, *motorcycle*, *bicycle*, *caravan*, *trailer*, *onrails*, *road*, *sidewalk*, *ground*, *parking*, *railtrack*, *polegroup*, *billboard*, *streetlight*, *building*, *wall*, *fence*, *guardrail*, *bridge*, *tunnel*, *vegetation*, *terrain*, *sky*, *unlabeled*, *outofroi*, *static*, and *dynamic* are described in the supplemental material to the Cityscapes [1] paper.

- The categories *motorcyclist*, *bicyclist*, *otherrider*, *othervehicle*, *wheeledslow*, *boat*, *roadmarking* (==*marking general*) *curb*, *bikelane*, *pedestrianarea*, *crosswalkplain*, *crosswalkzebra*, *servicelane*, *curbcut*, *water*, *sand*, *snow*, *pole*, *utilitypole*, *trafficlight*, *trafficsign* (== *traffic sign front*), *manhole*, *pothole*, *trafficsignback*, *trafficsignframe*, *otherbarrier*, *catchbasin*, *junctionbox*, *mailbox*, *phonebooth*, *bikerack*, *trashcan*, *bench*, *banner*, *firehydrant*, *cctvcamera*, *groundanimal*, *bird*, *mountain*, *dashcammount* (== *car mount*) are described in the supplemental material to the MVD [2] paper.

	human	vehicle	flat	object	construction	nature	sky	average
mvd100 PQ_{Cat}	46.0%	55.3%	71.6%	32.3%	52.8%	66.5%	79.5%	57.7%
mix150 PQ_{Cat}	49.7%	61.4%	86.2%	34.1%	62.5%	71.9%	87.3%	64.7%
mvd100 RQ_{Cat}	60.2%	67.0%	84.1%	48.2%	70.0%	82.1%	86.3%	71.1%
mix150 RQ_{Cat}	65.2%	73.6%	97.5%	50.9%	79.8%	87.1%	93.6%	78.2%
mvd100 SQ_{Cat}	76.4%	82.6%	85.1%	67.1%	75.4%	80.9%	92.2%	80.0%
mix150 SQ_{Cat}	76.2%	83.4%	88.4%	66.9%	78.2%	82.6%	93.3%	81.3%

Table 2. Per-supercategory PQ, RQ and SQ metrics evaluated on the hidden WD2 benchmark set for both models *mvd100* and *mix150* presented in the main paper.

- The labels *curb* and *curbterrain* both are described by the MVD *curb* label (i.e. curb stones; including all visible faces of a curb). If the curb encases an area labeled with terrain (or other vegetation), then the curb receives the *curbterrain* label. Otherwise use *curb*.
- The labels *bikelane* and *bikelanesidewalk* are both described by the MVD *curb* label. Use *bikelanesidewalk* if the bikelane is on a sidewalk. Otherwise use *bikelane*.
- The labels *manhole* and *manholesidewalk* are both described by the MVD *manhole* label. Use *manholesidewalk* if the manhole is on a sidewalk. Otherwise use *manhole*.
- The *traintrack* is described by the Cityscapes *traintrack* label (i.e. track of raised rails, not drivable by cars). The *tramtrack* label is used for the track area between embedded rails (drivable by cars) including the rails themselves.
- The *autorickshaw* category is described in the IDD [3] paper.
- The *trafficsignany* category is a fallback category used for cases where either *trafficsign* (=front) or *trafficsignback* could be correct.
- Vehicle class *pickup*: This label is used for light commercial vehicles (LCV) with an open cargo area. It only applies to motorized, car-sized vehicles with a visible un-roofed cargo area (also for open cages). Pickup trucks have regular car front wheels and a regular car wheelbase (distance). Cargo vehicles with larger tires or a truck motor housing (driver sitting above motor with a vertical windscreen and bonnet) retain the *truck* label.
- Vehicle class *van*: This label applies to motorized LCV without an open cargo area. Vans have a boxy shape with regular car tires and their wheel-base is typically larger than those of regular cars. The front of vans is often inclined but straight and they are distinctively higher (i.e. the van's ceiling) than regular cars. Vehicles sold under the term "mini-van" with a regular car height remain at the *car* label. Hybrid vehicles with a fully separated, non-continuous, driver cabin are still labeled as *truck* (e.g. many ambulance vehicles, police, some delivery trucks).

3. Supercategory Scores

Table 2 shows individual per-supercategory scores of mvd100 and mix150 on the $WD2_bench$ set. Table 3 in the main paper contains the right-most column (arithmetic mean over all supercategories) denoted as PQ_{cat}.

4. Further Experiments

The following Table 3 mirrors the layout of Table 3 in the main paper. Experiments were conducted for WD2, Cityscapes, and IDD in the same way as for MVD: first 100 epochs on only the original dataset using the standard training label policy (plus *van* and *pickup*. This results in 66 labels for MVD, 29 for IDD, and 22 for Cityscapes. The validation results are calculated on the original validation dataset using the original training label policy while the WD2 benchmark evaluation remaps the algorithm outputs to the WD2_{eval} label policy (26 labels) and reports the results for the hidden WD2 benchmark dataset. The models after 100 epochs are fine-tuned for 50 epochs using WD2 individually remapped to each of the dataset's training label policies mixed with the same amount of frames from the original dataset (i.e. 50%/50% split). The model for the first column *wd2_100* uses only WD2 frames with the WD2_{eval} label policy during training. The same train/val split is used both for *wd2_100* as well as in all fine-tunings.

	Original Validation					WD2 Benchmark						
	PQ	SQ	RQ	PQ_{van}	PQ_{pickup}	PQ	SQ	RQ	PQ_{van}	PQ_{pickup}	PQ_{neg}	PQ_{cat}
wd2_100	38.0%	75.6%	48.2%	36.9%	33.4%	37.0%	75.5%	47.7%	35.1%	37.3%	16.9%	61.1%
cs100	55.7%	76.4%	68.2%	53.7%	0.0%	10.7%	69.5%	15.0%	10.6%	0.0%	5.4%	22.8%
cs150	56.1%	77.4%	68.2%	55.2%	0.0%	32.2%	76.9%	41.1%	34.4%	41.1%	15.5%	58.4%
idd100	47.7%	75.6%	59.5%	48.8%	0.0%	15.3%	72.8%	20.1%	14.1%	0.0%	7.2%	35.7%
idd150	46.4%	75.7%	57.9%	40.4%	0.0%	29.4%	76.1%	37.7%	33.7%	33.7%	14.2%	54.8%
mvd100	35.1%	74.2%	43.9%	26.6%	29.9%	37.6%	75.6%	48.3%	34.0%	38.1%	17.1%	57.7%
mix150	34.1%	73.5%	42.8%	24.7%	29.7%	42.2%	77.5%	53.2%	38.9%	49.2%	21.1%	64.7%

Table 3. Comparison of performances for multiple models first trained on the respective original datasets (WD2, Cityscapes, IDD, MVD) and later fine-tuned for 50 additional epochs on a 50%/50% mixture of the original dataset and WD2. The left side shows results when evaluated on the original validation sets and the right side shows scores on the hidden WD2 benchmark set. The label policy on the left is not constant across the rows while the right side all evaluate using the same labels: WD2$_{eval}$. Results for MVD are duplicated from the main paper to improve comparability.

References

[1] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3213–3223, 2016. 1

[2] Gerhard Neuhold, Tobias Ollmann, Samuel Rota Bulo, and Peter Kontschieder. The mapillary vistas dataset for semantic understanding of street scenes. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4990–4999, 2017. 1

[3] Girish Varma, Anbumani Subramanian, Anoop Namboodiri, Manmohan Chandraker, and CV Jawahar. IDD: A dataset for exploring problems of autonomous navigation in unconstrained environments. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1743–1751. IEEE, 2019. 2