Supplementary Material
CLEO: Continual Learning of Evolving Ontologies

Shishir Muralidhara[1], Saqib Bukhari[2], Georg Schneider[2], Didier Stricker[1,3], and René Schuster[1,3]

[1] DFKI – German Research Center for Artificial Intelligence
[2] ZF Friedrichshafen AG
[3] RPTU -– University of Kaiserslautern-Landau
{firstname.lastname}@{dfki.de, zf.com}

1 Overview

This document provides the supplementary material to our main paper on *CLEO: Continual Learning of Evolving Ontologies*. We first present the detailed class hierarchies that we use to design our experiments for Cityscapes [2], PASCAL VOC [4], and Mapillary Vistas [5]. Afterwards, we extend our motivation for the design of these experiments, *i.e.* explain how they differ and which possible types of evolution are covered in each experiment. Additionally, we present the task-wise evaluation of our approach MoOn and the state-of-the-art for Cityscapes and provide a visualization for the task-wise evolution. The task-wise evaluation is extended by a detailed breakdown of per-class results after each task. Finally, we list the exact sets of classes for each task of each experiment as well as the final groups of classes that are considered in our evaluation of CLEO.

2 Class Hierarchies

2.1 Cityscapes

The Cityscapes dataset [2] provides annotations for 30 classes categorized into 8 groups namely, *flat, construction, object, nature, sky, human,* and *void* classes. Within these 30 classes, 19 classes are used for training and evaluation, while the remainder are treated as void. These 19 classes and their grouping into the 7 parent classes are visualized in Fig. 1. Notably, *sky* is the only class which in turn does not contain any further classes. The dataset exhibits varying levels of representation among the individual classes, with some being underrepresented, which significantly impacts the learning process. We use this class hierarchy to derive two settings for the Cityscapes dataset and list the classes included in each task in Tabs. 4 and 5.

2.2 PASCAL VOC 2012

The PASCAL VOC 2012 dataset [4] is widely used for benchmarking tasks such as classification, object detection, and semantic segmentation. It provides

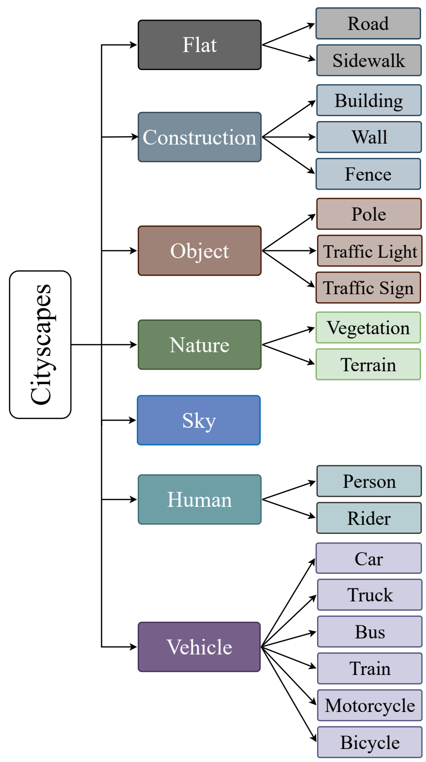

Fig. 1: The official label hierarchy of Cityscapes [2].

annotations for 20 object categories including common objects. A label hierarchy for PASCAL VOC is provided by [7], which groups these 20 classes into 4 parent classes with *person* being the only class without any sub-classes. Interestingly, the suggested hierarchy for PASCAL VOC consists of intermediate sub-groupings of these 20 classes into groups such as *domestic animals, farmyard animals, furniture, 4-wheeler,* and *2-wheeler*, resulting in a more challenging multi-level splitting and retention. The class hierarchy representing the grouping of the 20 original classes at different levels, including their sub-groups is presented in Fig. 2. The task-wise evolution for the experiments on PASCAL VOC is given in Tabs. 6 to 8.

2.3 Mapillary Vistas

The initial version v1.2 of the Mapillary Vistas dataset [5] released in 2017 provided annotations for 66 object classes, which was later refined into 124 classes in the current version v2.0. The Mapillary Vistas dataset does not provide an official mapping between the two versions. However, since both versions of the dataset annotate the same set of images, we can create a class mapping by comparing the ground truth of both versions to determine how the classes have

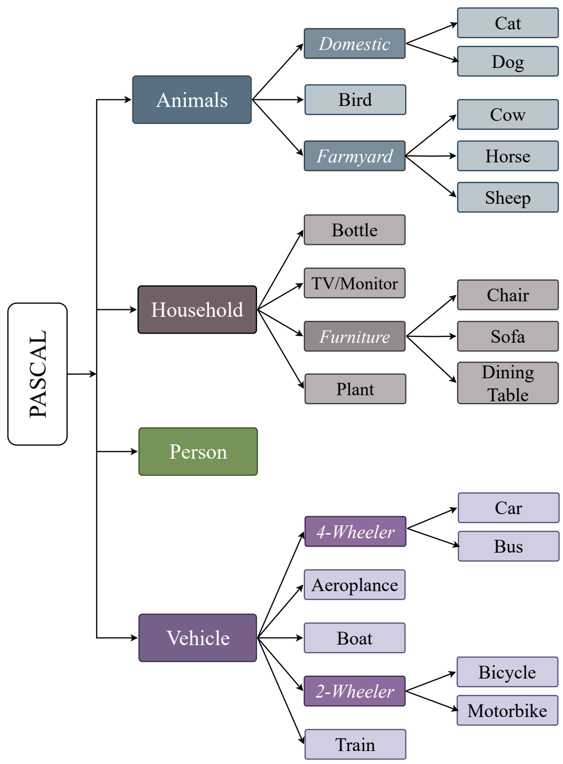

Fig. 2: Label hierarchy for PASCAL VOC [4] adapted from [7].

changed. The class mapping is visualized in Fig. 3. Through this process, we have identified ten parent classes in the initial version, from which the newly defined classes in v2.0 have originated. These classes include: *barrier, lane marking - crosswalk, parking, road, traffic sign (front), traffic sign (back), unlabelled, billboard, lane marking - general, and traffic light*. Version 2.0 introduces 14 entirely new classes that were previously grouped under the *unlabelled* class, while the remaining have been split from existing classes to represent more specific semantics. Notably, the classes, *billboard, lane marking - general, and traffic light* have been completely split into new classes and are no longer part of the current version, while the other parent classes have been split up only partially.

3 Differences between the Experiments

In this section, we describe the variations within our seven different experiments. The most prominent dimension along we vary is the dataset itself. This variation implies application domains, numbers of classes, different sensors, *etc*. Our experiments on Cityscapes [2] cover two potential cases. Firstly, in *CS-Ex1*, the semantic evolution affects multiple superclasses in parallel, *i.e.* during the same

Fig. 3: Class mapping for between both versions of Mapillary Vistas [5].

Table 1: Task-wise results in mIoU for CS-Ex1 after learning the final task.

Method	Task 0	Task 1	Task 2	Task 3	Task 4	Task 5	All
Fine-Tuning	00.00	00.00	00.00	00.00	00.00	00.00	00.00
Joint Training	47.31	73.84	45.65	64.03	48.76	41.08	55.62
MiB [1]	17.27	03.76	02.35	00.00	14.32	00.00	09.11
PLOP [3]	20.72	68.45	00.24	00.00	00.93	00.00	28.91
RCIL [6]	17.20	00.00	00.00	00.00	00.00	00.00	06.88
MoOn (Ours)	39.02	71.59	10.43	00.72	12.44	00.00	**39.31**

task. Secondly, within the same task, multiple subclasses are split from the same superclass in *CS-Ex2*. As such, these two experiments cover varying numbers of involved super- and subclasses.

The applied label hierarchy of PASCAL VOC [4] adds an additional level of semantic resolution (*cf*. Fig. 2), which allows us to design experiments in which splitting and retention can happen more than once, *i.e.* a single class can undergo two evolutionary steps.

Lastly, with the vast number of classes in Mapillary Vistas [5], we can cover two more use-cases and can define more challenging and complex scenarios. *I.e.*, in *MV-Ex1* we cover the realistic semantic evolution that was introduced by the authors of the dataset, that affects many super- and subclasses in a single task. Due to the large number of classes, we can also design an experiment that covers a much longer sequence of tasks in *MV-Ex2*, *i.e.* 10 evolutionary steps.

4 Task-wise Results

Table 1 shows task-wise results on Cityscapes after training on the final task. We have computed the mIoU for the classes of each task and compare our approach MoOn to MiB [1], PLOP [3], and RCIL [6]. Fine-tuning typically leads to catastrophic forgetting of previous tasks, and only information of classes learned in the final task is retained. However for *CS-Ex1* not even this is the case, since the last task contains the highly under-represented class *motorcycle*. All CL approaches struggle with task 3 and 5 In this experiment for the same reason. MoOn demonstrates the lowest amount of forgetting of the initial set of classes in task 0 and the highest plasticity when learning new classes throughout the evolution of the ontology. Figure 4 visualizes the results of MoOn for three validation samples of Cityscapes to show how the semantic ontology evolves over time in *CS-Ex1*. The figure visualizes the 19 classes that are grouped into 7 parent classes in task 0. In task 1, the road class is split from the parent class *flat*, retaining the *sidewalk* class, and the car class is split from the parent class *vehicle*. The *vehicle* class continues to split in subsequent tasks, such as into the *bus* class in task 3.

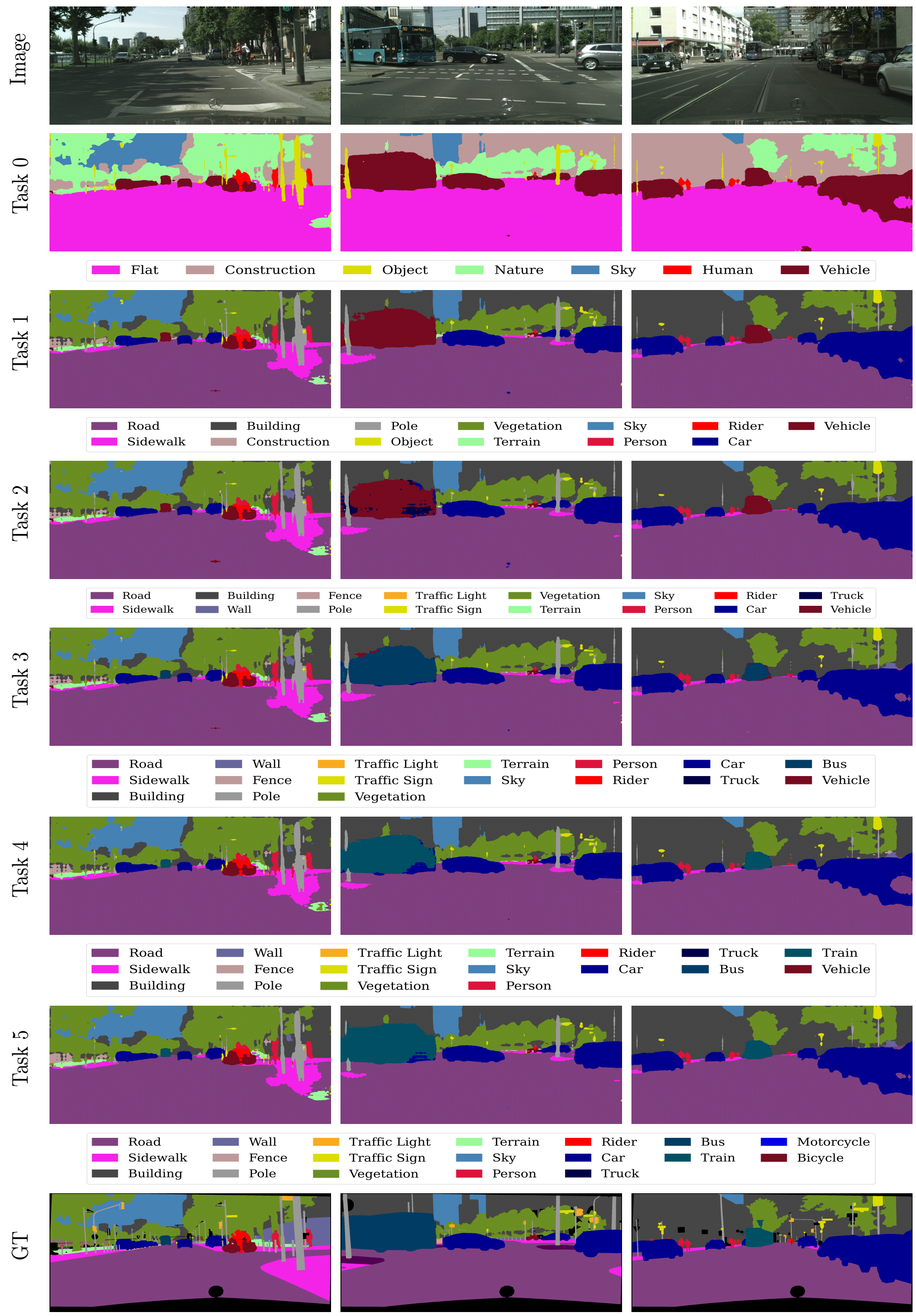

Fig. 4: Task-wise visualization for CLEO setting *CS-Ex1* on the Cityscapes dataset [2]. The sequential refinement of classes can be observed at each evolutionary step.

5 Class-wise Results

Tables 2 and 3 show the class-wise results of our approach MoOn after each task. The results shed light on the influence of the splitting order. In case of *CS-Ex1*, one class is split from each of the parent classes whereas in *CS-Ex2*, each parent class is split completely in each step. We can observe that in *CS-Ex2*, the results for the *Vehicle* subclasses are better compared to *CS-Ex1*, as it splits completely in a single final step. On the contrary, we observe that a few classes which are split in the earlier tasks achieve better results in *CS-Ex1*, *e.g.* the *wall* or *rider* classes.

Table 2: Class-wise results for CS-Ex1 after learning each task.

	road	*sidewalk*	*building*	*wall*	*fence*	*pole*	*traffic light*	*traffic sign*	*vegetation*	*terrain*	*sky*	*person*	*rider*	*car*	*truck*	*bus*	*train*	*motorcycle*	*bicycle*
Task	**Flat**		**Construction**			**Object**			**Nature**		**Sky**	**Human**		**Vehicle**					
0	97.99		86.78			33.74			88.58		90.40	65.65		85.01					
1	96.95	59.11	86.44	36.94		19.49	42.29		88.15	41.61	90.41	61.65	15.23	83.59	58.90				
2	96.93	59.18	86.32	31.92	22.47	19.35	00.00	39.19	88.13	41.76	90.29	61.47	14.89	82.78	23.69	50.90			
3	96.88	58.91	86.26	32.06	21.42	17.90	00.00	38.34	88.09	41.81	90.21	61.24	14.68	83.27	00.00	37.73	44.38		
4	96.84	58.56	86.19	28.10	20.21	16.15	00.00	37.48	87.84	41.69	90.17	60.63	15.55	82.10	00.00	1.42	13.11	50.55	
5	96.79	58.17	86.15	31.31	21.72	15.65	00.00	37.92	88.00	41.81	90.21	60.39	13.01	82.59	00.00	00.72	12.44	00.00	49.38

Table 3: Class-wise results for CS-Ex2 after learning each task.

	road	*sidewalk*	*building*	*wall*	*fence*	*pole*	*traffic light*	*traffic sign*	*vegetation*	*terrain*	*sky*	*person*	*rider*	*car*	*truck*	*bus*	*train*	*motorcycle*	*bicycle*
Task	**Flat**		**Construction**			**Object**			**Nature**		**Sky**	**Human**		**Vehicle**					
0	97.99		86.78			33.74			88.58		90.40	65.65		85.01					
1	96.99	59.25	86.76			33.79			88.56		90.39	65.49		84.97					
2	96.98	59.05	86.51	24.21	21.91	34.01			88.54		90.41	65.35		84.72					
3	96.98	59.01	86.38	24.17	22.33	20.23	00.00	37.2	88.42		90.29	65.01		84.67					
4	96.95	58.65	86.38	24.65	21.98	19.38	00.00	37.35	88.18	41.05	90.32	64.88		84.55					
5	96.93	58.35	86.29	25.59	21.44	18.73	00.00	37.22	88.18	40.88	90.29	61.24	02.31	84.44					
6	96.89	58.36	86.13	26.40	22.12	18.38	00.00	37.52	88.17	40.55	90.24	60.97	06.15	80.47	01.64	25.96	22.44	00.00	50.00

6 Task-wise Semantic Evolution

For maximum transparency, we explicitly list the classes C_t for each task of each of our experiments. For the two experimental setting on Cityscapes, the evolution of classes is described in Tabs. 4 and 5, for PASCAL VOC, the evolution is given in Tabs. 6 to 8, and for Mapillary Vistas, we list all evolutionary steps in Tabs. 9 and 10. Additionally, in each table, we list all the classes contained in each evaluation category (*unsplit*, *split*, *retained*) after the sequential learning.

Table 4: Evolutionary steps for CS-Ex1.

CS-Ex1	
Task t	C_t
0	*background, flat, construction, object, nature, sky, human, vehicle*
1	*road, building, pole, vegetation, person, car*
2	*wall, traffic light, truck*
3	*bus*
4	*train*
5	*motorcycle*
Class Groups	
Unsplit	*sky*
Split	*road, building, pole, vegetation, person, car, wall, traffic light, truck, bus, train, motorcycle*
Retained	*flat (sidewalk), construction (fence), object (traffic sign), nature (terrain), human (rider), vehicle (bicycle)*

Table 5: Evolutionary steps for CS-Ex2.

CS-Ex2	
Task t	C_t
0	*background, flat, construction, object, nature, sky, human, vehicle*
1	*road*
2	*building, wall*
3	*pole, traffic light*
4	*vegetation*
5	*person*
6	*car, truck, bus, train, motorcycle*
Class Groups	
Unsplit	*sky*
Split	*road, building, wall, pole, traffic light, vegetation, person, car, truck, bus, train, motorcycle*
Retained	*flat (sidewalk), construction (fence), object (traffic sign), nature (terrain), human (rider), vehicle (bicycle)*

Table 6: Evolutionary steps for VOC-Ex1.

VOC-Ex1	
Task t	C_t
0	*background, animals, household, person, vehicle*
1	*farmyard, bird, bottle, furniture, 2-wheeler, aeroplane*
2	*cow, horse, sheep, chair, sofa, dining table, bicycle, motorbike*
Class Groups	
Unsplit	*person*
Split	*bird, bottle, aeroplane, cow, horse, sheep, chair, sofa, dining table, bicycle, motorbike*
Retained	*animals (cat, dog), household (tv/monitor, plant), 4-wheeler (car, bus), vehicle (boat, train)*

Table 7: Evolutionary steps for VOC-Ex2.

VOC-Ex2	
Task t	C_t
0	*background, animals, household, person, vehicle*
1	*bird, plant, train*
2	*sheep, tv/monitor, boat*
3	*horse, dining table, aeroplane*
4	*cow, sofa, motorbike*
5	*dog, chair, bicycle*
Class Groups	
Unsplit	*person*
Split	*bird, plant, train, sheep, tv/monitor, boat, horse, dining table, aeroplane, cow, sofa, motorbike, dog, chair, bicycle*
Retained	*animals (cat), household (bottle), vehicle (car, bus)*

Table 8: Evolutionary steps for VOC-Ex3.

VOC-Ex3	
Task t	C_t
0	*background, animals, household, person, vehicle*
1	*dog, horse, cow, sheep, bird*
2	*chair, sofa, table, tv/monitor, plant*
3	*bus, bicycle, motorbike, aeroplane, boat, train*
Class Groups	
Unsplit	*person*
Split	*dog, horse, cow, sheep, bird, chair, sofa, table, tv/monitor, plant, bus, bicycle, motorbike, aeroplane, boat, train*
Retained	*animals (cat), household (bottle), vehicle (car)*

Table 9: Evolutionary steps for MV-Ex1.

MV-Ex1	
Task t	C_t
0	*background, bird, ground animal, curb, fence, guard rail, barrier, wall, bike lane, crosswalk - plain, curb cut, parking, pedestrian area, rail track, road, service lane, sidewalk, bridge, building, tunnel, person, bicyclist, motorcyclist, other rider, lane marking - crosswalk, lane marking - general, mountain, sand, sky, snow, terrain, vegetation, water, banner, bench, bike rack, billboard, catch basin, cctv camera, fire hydrant, junction box, mailbox, manhole, phone booth, pothole, street light, pole, traffic sign frame, utility pole, traffic light, traffic sign (back), traffic sign (front), trash can, bicycle, boat, bus, car, caravan, motorcycle, on rails, other vehicle, trailer, truck, wheeled slow, car mount, ego vehicle*
1	*ambiguous barrier, concrete block, driveway, dynamic, garage, ground, lane marking (only) - crosswalk, lane marking (only) - dashed line, lane marking (only) - other, lane marking (only) - test, lane marking - ambiguous, lane marking - arrow (left), lane marking - arrow (other), lane marking - arrow (right), lane marking - arrow (split left or straight), lane marking - arrow (split right or straight), lane marking - arrow (straight), lane marking - give way (row), lane marking - give way (single), lane marking - hatched (chevron), lane marking - hatched (diagonal), lane marking - other, lane marking - stop line, lane marking - straight line, lane marking - symbol (bicycle), lane marking - symbol (other), lane marking - text, lane marking - zigzag line, lane separator, parking aisle, parking meter, person group, pole group, road median, road shoulder, road side, signage - ambiguous, signage - back, signage - information, signage - other, signage - store, static, temporary barrier, traffic cone, traffic island, traffic light - cyclists, traffic light - general (horizontal), traffic light - general (upright), traffic light - other, traffic light - pedestrians, traffic sign - ambiguous, traffic sign - direction (back), traffic sign - direction (front), traffic sign - parking, traffic sign - temporary (back), traffic sign - temporary (front), vehicle group, water valve*

Class Groups	
Unsplit	*bird, ground animal, curb, fence, guard rail, wall, bike lane, crosswalk - plain, curb cut, pedestrian area, rail track, service lane, sidewalk, bridge, building, tunnel, person, bicyclist, motorcyclist, other rider, mountain, sand, sky, snow, terrain, vegetation, water, banner, bench, bike rack, catch basin, cctv camera, fire hydrant, junction box, mailbox, manhole, phone booth, pothole, street light, pole, traffic sign frame, utility pole, trash can, bicycle, boat, bus, car, caravan, motorcycle, on rails, other vehicle, trailer, truck, wheeled slow, car mount, ego vehicle*
Split	*ambiguous barrier, concrete block, driveway, dynamic, garage, ground, lane marking (only) - crosswalk, lane marking (only) - dashed line, lane marking (only) - other, lane marking (only) - test, lane marking - ambiguous, lane marking - arrow (left), lane marking - arrow (other), lane marking - arrow (right), lane marking - arrow (split left or straight), lane marking - arrow (split right or straight), lane marking - arrow (straight), lane marking - give way (row), lane marking - give way (single), lane marking - hatched (chevron), lane marking - hatched (diagonal), lane marking - other, lane marking - stop line, lane marking - straight line, lane marking - symbol (bicycle), lane marking - symbol (other), lane marking - text, lane marking - zigzag line, lane separator, parking aisle, parking meter, person group, pole group, road median, road shoulder, road side, signage - ambiguous, signage - back, signage - information, signage - other, signage - store, static, temporary barrier, traffic cone, traffic island, traffic light - cyclists, traffic light - general (horizontal), traffic light - general (upright), traffic light - other, traffic light - pedestrians, traffic sign - ambiguous, traffic sign - direction (back), traffic sign - direction (front), traffic sign - parking, traffic sign - temporary (back), traffic sign - temporary (front), vehicle group, water valve*
Retained	*background, barrier, lane marking - crosswalk, parking, road, traffic sign (front), traffic sign (back), signage - advertisement, lane marking - dashed line, traffic light - general (single)*

Table 10: Evolutionary steps for MV-Ex2.

MV-Ex2	
Task t	C_t
0	*background, bird, ground animal, curb, fence, guard rail, barrier, wall, bike lane, crosswalk - plain, curb cut, parking, pedestrian area, rail track, road, service lane, sidewalk, bridge, building, tunnel, person, bicyclist, motorcyclist, other rider, lane marking - crosswalk, lane marking - general, mountain, sand, sky, snow, terrain, vegetation, water, banner, bench, bike rack, billboard, catch basin, cctv camera, fire hydrant, junction box, mailbox, manhole, phone booth, pothole, street light, pole, traffic sign frame, utility pole, traffic light, traffic sign (back), traffic sign (front), trash can, bicycle, boat, bus, car, caravan, motorcycle, on rails, other vehicle, trailer, truck, wheeled slow, car mount, ego vehicle*
1	*concrete block, road median, road side, lane separator*
2	*lane marking (only) - crosswalk*
3	*parking aisle*
4	*road shoulder*
5	*temporary barrier, traffic sign - direction (front), traffic sign - parking, traffic sign - temporary (front)*
6	*traffic sign - direction (back), traffic sign - temporary (back)*
7	*ambiguous barrier, driveway, traffic island, garage, person group, parking meter, pole group, traffic cone, traffic sign - ambiguous, vehicle group, water valve, dynamic, ground, static*
8	*signage - ambiguous, signage - back, signage - information, signage - other, signage - store*
9	*lane marking - straight line, lane marking - zigzag line, lane marking - ambiguous, lane marking - arrow (left), lane marking - arrow (other), lane marking - arrow (right), lane marking - arrow (split left or straight), lane marking - arrow (split right or straight), lane marking - arrow (straight), lane marking - give way (row), lane marking - give way (single), lane marking - hatched (chevron), lane marking - hatched (diagonal), lane marking - other, lane marking - stop line, lane marking - symbol (bicycle), lane marking - symbol (other), lane marking - text, lane marking (only) - dashed line, lane marking (only) - other, lane marking (only) - test*
10	*traffic light - pedestrians, traffic light - general (upright), traffic light - general (horizontal), traffic light - cyclists, traffic light - other*

Class	Groups
Unsplit	*bird, ground animal, curb, fence, guard rail, wall, bike lane, crosswalk - plain, curb cut, pedestrian area, rail track, service lane, sidewalk, bridge, building, tunnel, person, bicyclist, motorcyclist, other rider, mountain, sand, sky, snow, terrain, vegetation, water, banner, bench, bike rack, catch basin, cctv camera, fire hydrant, junction box, mailbox, manhole, phone booth, pothole, street light, pole, traffic sign frame, utility pole, trash can, bicycle, boat, bus, car, caravan, motorcycle, on rails, other vehicle, trailer, truck, wheeled slow, car mount, ego vehicle*
Split	*ambiguous barrier, concrete block, driveway, dynamic, garage, ground, lane marking (only) - crosswalk, lane marking (only) - dashed line, lane marking (only) - other, lane marking (only) - test, lane marking - ambiguous, lane marking - arrow (left), lane marking - arrow (other), lane marking - arrow (right), lane marking - arrow (split left or straight), lane marking - arrow (split right or straight), lane marking - arrow (straight), lane marking - give way (row), lane marking - give way (single), lane marking - hatched (chevron), lane marking - hatched (diagonal), lane marking - other, lane marking - stop line, lane marking - straight line, lane marking - symbol (bicycle), lane marking - symbol (other), lane marking - text, lane marking - zigzag line, lane separator, parking aisle, parking meter, person group, pole group, road median, road shoulder, road side, signage - ambiguous, signage - back, signage - information, signage - other, signage - store, static, temporary barrier, traffic cone, traffic island, traffic light - cyclists, traffic light - general (horizontal), traffic light - general (upright), traffic light - other, traffic light - pedestrians, traffic sign - ambiguous, traffic sign - direction (back), traffic sign - direction (front), traffic sign - parking, traffic sign - temporary (back), traffic sign - temporary (front), vehicle group, water valve*
Retained	*background, barrier, lane marking - crosswalk, parking, road, traffic sign (front), traffic sign (back), signage - advertisement, lane marking - dashed line, traffic light - general (single)*

References

1. Cermelli, F., Mancini, M., Bulo, S.R., Ricci, E., Caputo, B.: Modeling the background for incremental learning in semantic segmentation. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2020)
2. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B.: The cityscapes dataset for semantic urban scene understanding. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2016)

3. Douillard, A., Chen, Y., Dapogny, A., Cord, M.: Plop: Learning without forgetting for continual semantic segmentation. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2021)
4. Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International Journal of Computer Vision (IJCV) (2010)
5. Neuhold, G., Ollmann, T., Rota Bulo, S., Kontschieder, P.: The mapillary vistas dataset for semantic understanding of street scenes. In: International Conference on Computer Vision (ICCV) (2017)
6. Zhang, C.B., Xiao, J.W., Liu, X., Chen, Y.C., Cheng, M.M.: Representation compensation networks for continual semantic segmentation. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2022)
7. Zhao, H., Hu, Q., Zhu, P., Wang, Y., Wang, P.: A recursive regularization based feature selection framework for hierarchical classification. Transactions on Knowledge and Data Engineering (2021)