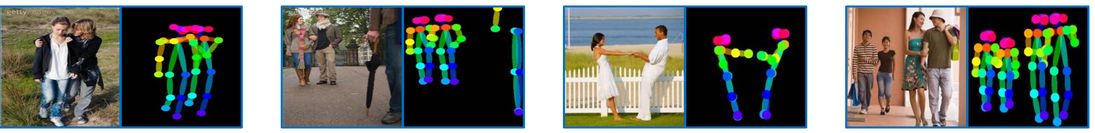
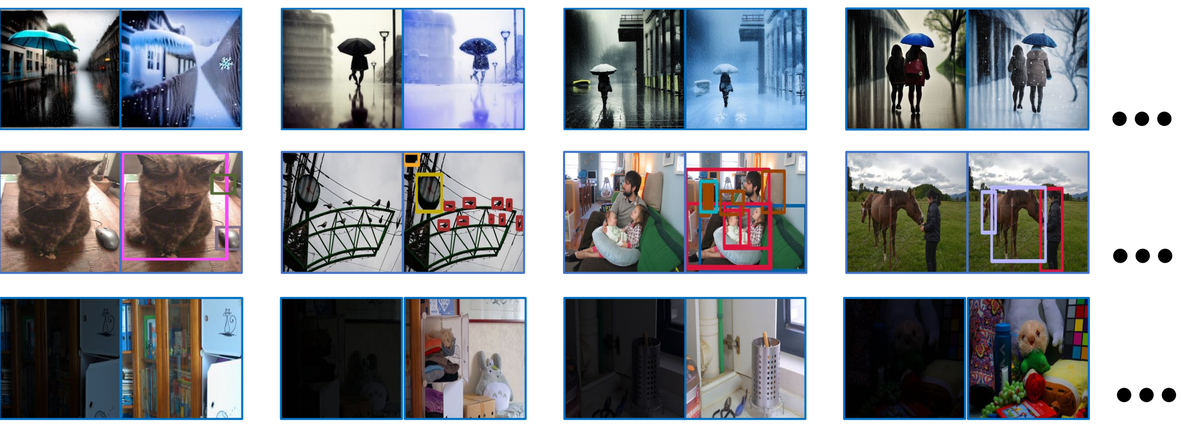
Visual Sentences

Single images, e.g. LAION

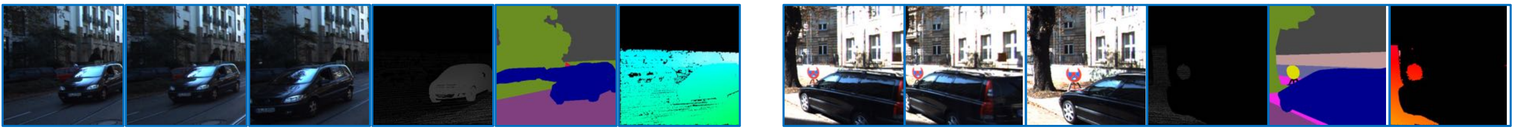
Image sequences, e.g. videos, 3D rotations, synthetic viewpoints



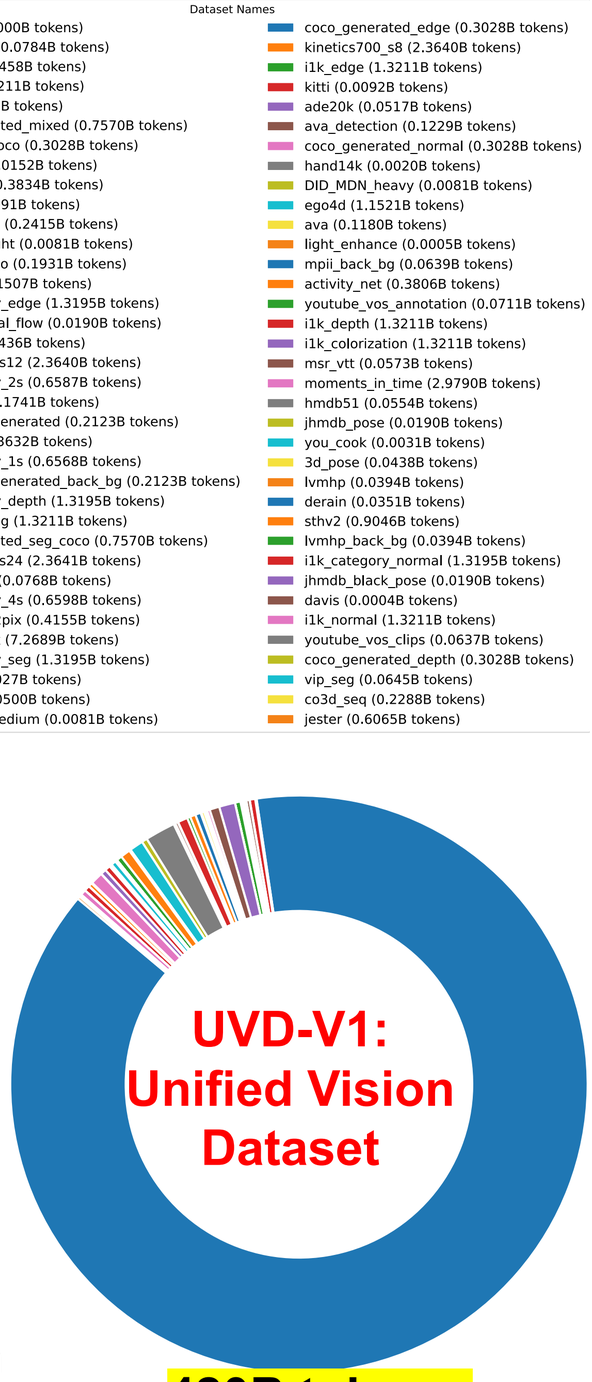


Images with annotation, e.g. style transfer, object detection, low light enhancement

Images with free form annotation, e.g. object detection + instance segmentation etc



Videos with annotation, e.g. video segmentation



420B tokens, 50 Datasets.