KI DELTA LEARNING
Scalable AI for Automated Driving

Unsupervised Semantic Image Synthesis

George Eskandar, Bin Yang

WP 2.2
WP 2.4
WP 2.5

Use-Case in Delta Learning

Semantic image synthesis (SIS) is the task of generating high resolution images from user-specified semantic layouts. SIS opens the door to an extensive range of applications such as content creation and semantic manipulation by editing, adding, removing or changing the appearance of an object. In autonomous driving, SIS can be used as a data augmentation tool that can rival driving simulators like CARLA as it allows a controlled synthesis of images while producing photorealistic results which do not suffer from a domain gap like synthetic images.

Technical Problem

The problem of semantic image synthesis has mostly been addressed as a supervised learning problem in the literature. Although state-of-the-art methods can produce visually appealing high-resolution images, they still depend on a lot of annotated paired data which is expensive to acquire.

Technical Solution

We propose an unsupervised framework which can synthesize realistic images from labels without the use of paired data.
An unsupervised paradigm for SIS is introduced, which involves an adversarial training between a generator and a whole image wavelet-based discriminator, and a cooperative training between the generator and a Unet. We provide the wavelet decomposition of the real and fake images as input to our discriminator so that it makes a decision based on higher spatial frequencies in the image.

Evaluation

We conduct extensive experiments on 3 image datasets (COCO-stuff, Cityscapes and ADE20K in an unpaired setting) to showcase the ability of our model to generate a high diversity of photorealistic images and close the gap between supervised and unsupervised methods in SIS..

Method	Supervised	Cityscapes		ADE20K		COCO-stuff	
		FID↓	mIoU↑	FID↓	mIoU↑	FID↓	mIoU↑
CycleGAN [83]	✗	87.2	24.5	94.7	8.5	126.2	3.5
MUNIT [35]	✗	84	8.2	n/a	n/a	n/a	n/a
DRIT [43]	✗	164	9.5	132.2	0.016	134.5	0.008
DistanceGAN [4]	✗	78	17.6	80	0.035	92.4	0.014
GCGAN [22]	✗	80	8.4	92	0.07	99.8	0.019
CUT [52]	✗	55.7	29.8	75.1	4.4	153.8	2.4
U-SIS	✗	**52.2**	**42.8**	**32.4**	**17.1**	**30.6**	**13.5**
CRN [10]	✓	73.3	22.4	104.7	52.4	70.4	23.7
SIMS [55]	✓	n/a	n/a	49.7	47.2	n/a	n/a
Pix2pixHD [72]	✓	81.8	20.3	95.0	58.3	111.5	14.6
SPADE [53]	✓	33.9	38.5	71.8	62.3	22.6	37.4

Table 1: Qualitative Results on 3 Datasets

Figure 1: Qualitative Results on Cityscapes dataset. We compare with CUT (unsupervised) and OASIS (supervised)

Environment

Universität Stuttgart

For more information contact:
george.eskandar@iss.uni-stuttgart.de

Partners

External partners

KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Energy.

KI FAMILIE

Supported by:
Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag

www.ki-deltalearning.de @KI_Familie KI Familie