CyEDA: CYCLE-OBJECT EDGE CONSISTENCY DOMAIN ADAPTATION

Paper ID 2886

Jing Chong Beh [1] Kam Woh Ng [2] Jie Long Kew [1] Che-Tsung Lin [3] Chee Seng Chan [1] Shang-Hong Lai [4, 5] Christopher Zach [3]

[1] CISiP, Faculty of Comp. Sci. and Info. Tech., Universiti Malaya, Kuala Lumpur, Malaysia [2] CVSSP, University of Surrey, Guildford, U.K.
[3] Dept. of Electrical Engineering, Chalmers University of Technology, Sweden [4] Microsoft AI R&D Center, Taiwan [5] Dept. of Computer Science, National Tsing Hua University, Taiwan

Introduction

We propose a new approach to achieve instance-level domain adaptation results without any detection subnet integration

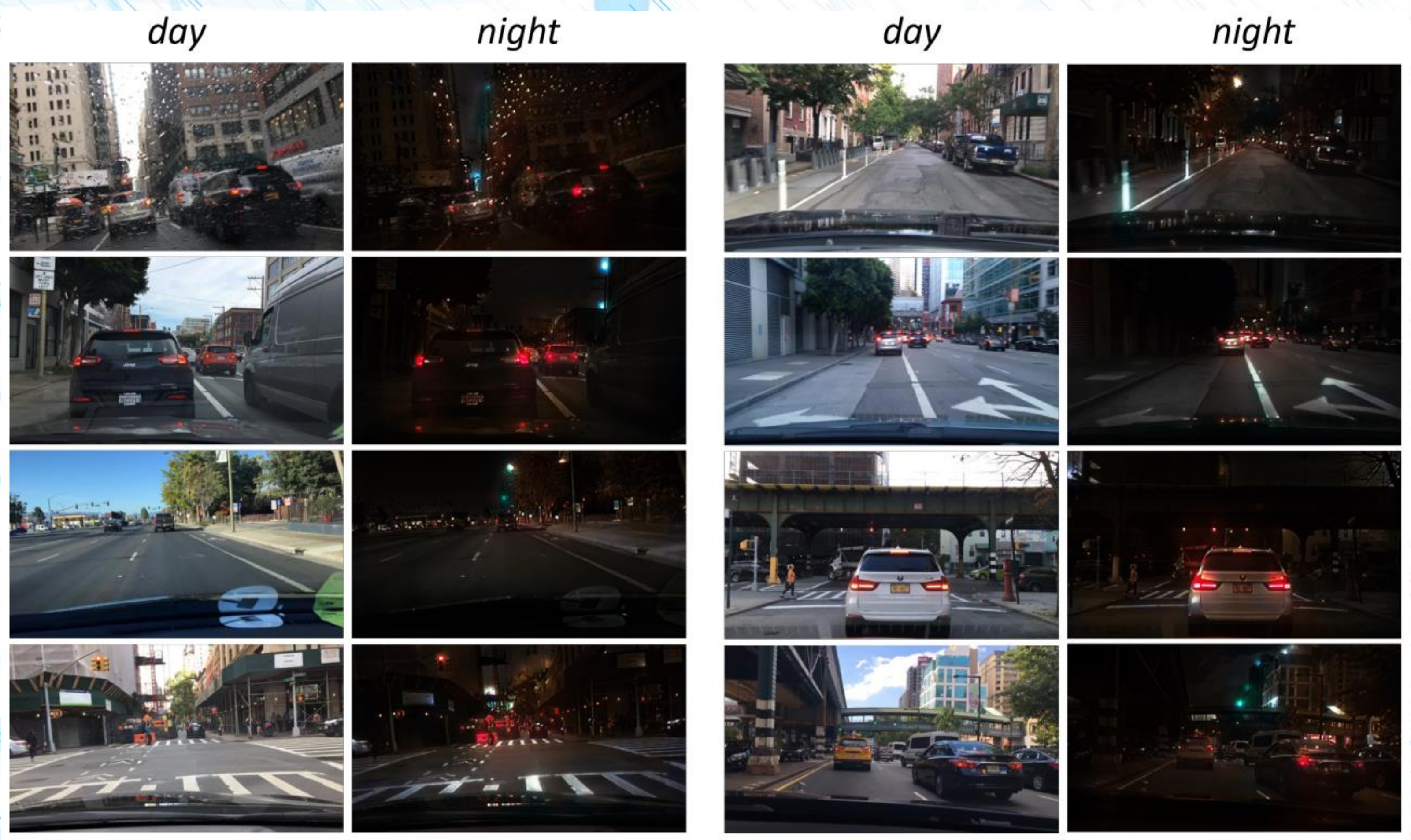
Our contributions consist of following:

- **Masking**
- **Cycle-Object Edge Consistency Loss**

Github repository link is available in the QR code.

Results

BDD100k day to night



SOTA Comparison

Qualitative & quantitative (FID score) comparison

GAN model	annotation?	GTA	BDD100k
MUNIT [16]	No	1.066	2.461
NICE-GAN [15]	No	2.466	1.913
AugGAN [5]	Yes	0.825	0.332
Multimodal AugGAN [6]	Yes	1.023	0.496
CyEDA (This work)	No	**0.737**	**0.297**

Methodology

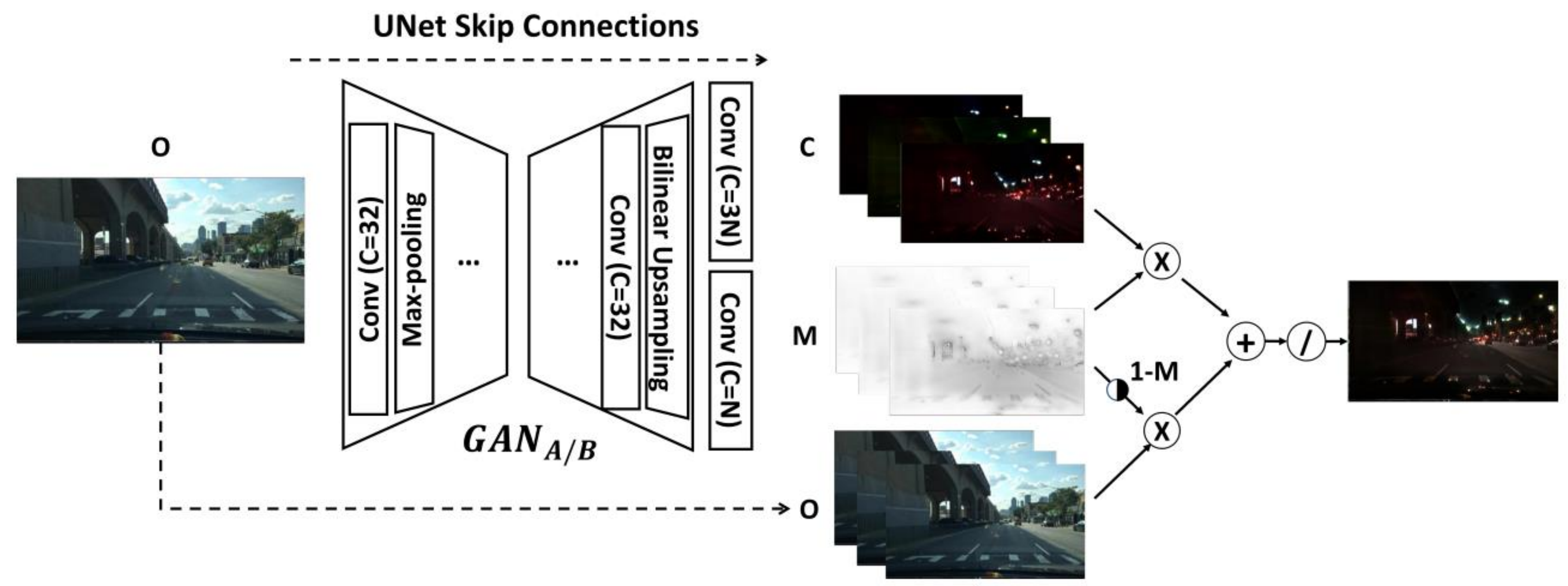
Mask UNet

Model learns to **make color changes to every pixel of original image** instead of generating whole image from hidden embedding.

We modify the last layer of UNet into two separate blocks:

- a convolution layer that yields *3N* channels of output followed by *tanh* activation layer (**predicted color changes**, *C*)
- a convolution layer that yields *N* channels of output followed by *sigmoid* activation layer (**predicted degree of color changes** / **mask**, *M*)

Translated image is computed by summing the weighted mask of color changes and the inverted weighted mask of the original image and normalized by number of mask *N*

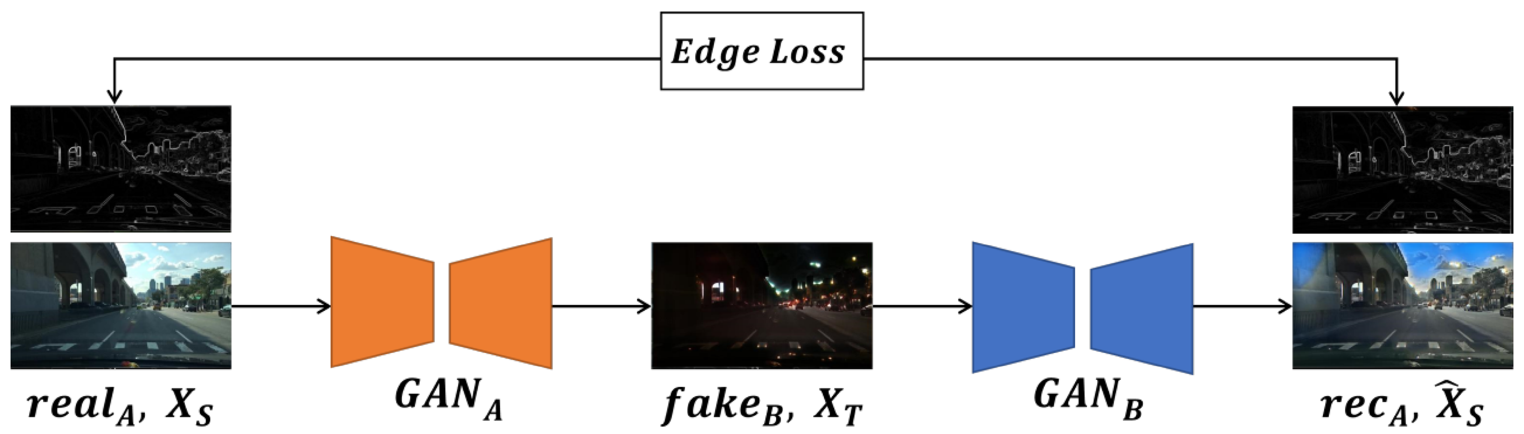


Cycle-Object Edge Consistency Loss

CycleGAN learns to hide information in the network to satisfy *L1* cycle consistency requirement [1].

We argue that the cycle consistency loss should only enforce the preservation of objects in the images instead of every pixel details.

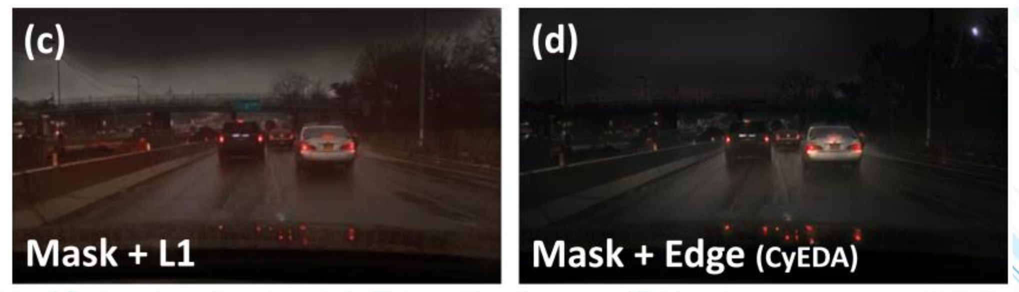
We **enforce *L1* consistency between edge of real and reconstructed images** instead of the image itself.



Ablation Study

Mask UNet contributes on **maintaining colour contrast** of translated image as can be seen in Fig.(c) and edge loss **removes unnecessary details** as shown in Fig.(d).

Qualitative and quantitative results of ablation study to show contributions of Mask Unet and Cycle-Object Edge Consistency Loss



Experiment Settings	FID
No mask + L1 (CycleGAN + UNet)	0.551
No mask + Edge	0.579
Mask + L1	0.389
Mask + Edge (CyEDA)	**0.297**

Domain Adaptation

YOLOv5s model is used to train with BDD100k real night images with and without translated day-to-night images.

Training Dataset (BDD100k)	mAP (whole)	AP (car)
det-train-night	0.444	0.627
+ det-val-day day-to-night	**0.465**	**0.644**

Conclusion

This paper introduced an approach to retain instance-level detail when translating images to a target domain by

- generating masks from UNet and performing color fine-tuning on original images according to the masks
- using cycle-object edge consistency loss to remove unnecessary details and provide extra capacity for model to perform more realistic image-translation

Reference

1. Casey Chu, Andrey Zhmoginov, and Mark Sandler, "Cyclegan, a master of steganography," arXiv preprint arXiv:1712.02950, 2017.

UNIVERSITI MALAYA UNIVERSITY OF SURREY CHALMERS UNIVERSITY OF TECHNOLOGY Microsoft 國立清華大學 NATIONAL TSING HUA UNIVERSITY ICIP2022 BORDEAUX