首届致远学术节
学生科研成果展示

Virtual to Real Reinforcement Learning for Autonomous Driving

Xinlei Pan*[1], Yurong You*[2], Ziyan Wang[3], Cewu Lu[2]

[1] University of California, Berkeley, [2] Shanghai Jiao Tong University, [3] Tsinghua University, * indicates equal contribution

Spotlight
BMVC 2017

Motivation & Introduction

Reinforcement learning is considered as a promising direction for driving policy learning. However,

- **Training autonomous driving vehicles with reinforcement learning in real environment involves non-affordable trial-and-error.**
- One can first train the model in a virtual environment and then apply it in the real world. But **gaps exist between the real and virtual environment, rendering the transferring intractable.**

In this paper, we propose a novel **realistic translation network** to make model trained in virtual environment **more adaptable to real world scenario.** The proposed framework exploits a key relation between non-realistic virtual images and real images: **they all have similar scene structure**. Though lacking proper paired data labeling virtual images with corresponding real images, their segmentation can serve as a proxy: we can first convert the virtual images into their scene segmentations, then convert the segmentations into real images, where paired training data are available in both part.

Experiments show that our proposed virtual to real (VR) reinforcement learning (RL) works pretty well. To our knowledge, this is the first successful case of driving policy trained by reinforcement learning that can adapt to real world driving data.

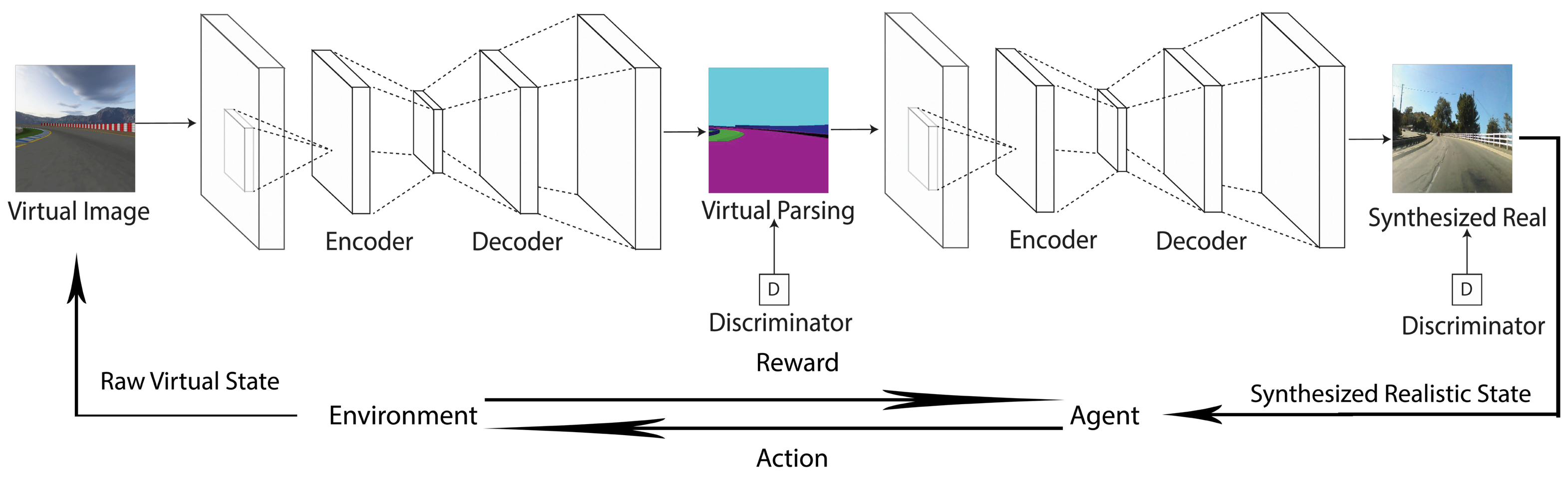


Figure 1: Framework for virtual to real reinforcement learning (VRRL) for autonomous driving

Methods

The overall pipeline is shown in *Figure 1*.
Inspired by [1], our realistic translation network is composed of two image translation networks:
1.**virtual-to-segmentation network**: translating virtual images to their segmentations
2.**segmentation-to-real network**: translating segmented images to their real world counterparts

These two networks are basically conditional GANs. The objective of a conditional GAN can be expressed as,

$$\begin{aligned}\mathcal{L}_{cGAN}(G,D) =& \mathbb{E}_{x,s\sim p_{data}(x,s)}[\log D(x,s)]\\ &+\mathbb{E}_{x\sim p_{data}(x),z\sim p_z(z)}[\log(1-D(x,G(x,z)))],\end{aligned}$$

where G is the generator that tries to minimize this objective and D is the adversarial discriminator that acts against G to maximize this objective. In order to suppress blurring, a L1 loss regularization term is added. Therefore, the overall objective for the image-to-image translation network is,

$$G^* = \arg\min_G\max_D \mathcal{L}_{cGAN}(G,D)+\lambda\mathcal{L}_{L1}(G),$$

where λ is the weight of regularization.

These two networks are both trained towards the above defined generator, where first network translates virtual images x to their segmentations $s : G_1 : \{x,z_1\} \rightarrow s$, and the second network translates segmented images s into their realistic counterparts $y : G_2 : \{s,z_2\} \rightarrow y$, where z_1 ,z_2 are noise terms to avoid deterministic outputs. We use the same generator and discriminator architectures as used in [1].

After training of the realistic translation network is finished, we train our **reinforcement learning model** by using it to filter virtual images to synthetic realistic images and feed these realistic images as state inputs.

For more details about the proposed architecture, including network settings of translation networks and reinforcement learning models, please refer to our paper.

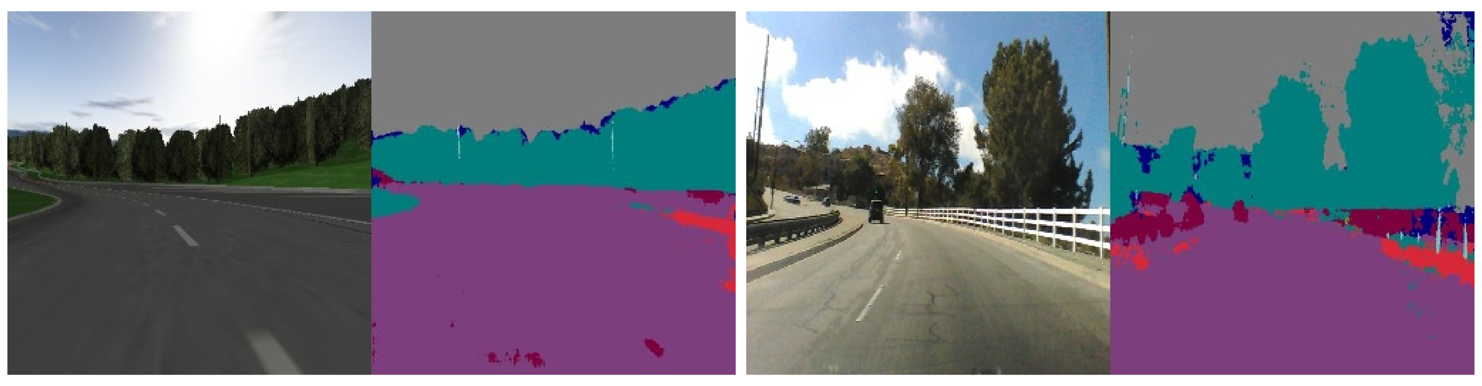


Figure 2: Data examples. Left: segmentation for a virtual image, Right: segmentation for a real image

Data

Data for **Segmentation-to-real network**:
The real world driving video data are from [2], which is collected in a sunny day with detailed steering angle annotations per frame .We used the image semantic segmentation network design of [3] and their trained segmentation network on the CityScape image segmentation dataset to segment 45k real world driving images from [2]. The network was trained on the CityScape dataset with 11 classes and was trained with 30000 iterations.

Data for **Segmentation-to-real network:**
We collected virtual images and their segmentations from the Aalborg environment in TORCS [4]. A total of 1673 images were collected which covers the entire driving cycle of Aalborg environment.

Figure 3: Examples of Virtual to Real Image Translation.
Odd Columns: virtual images, Even Columns: synthetic real world images

Results

We performed two sets of experiments to compare the performance of our method and other reinforcement learning methods as well as supervised learning methods:
1.virtual to real reinforcement learning on real world driving data.
2.transfer learning in different virtual driving environments.
The virtual simulator used in our experiments is TORCS [4].

Qualitative Result of Realistic Translation Network

Representative results of our image translation network are shown in *Figure. 2*. The translation quality is satisfactory.

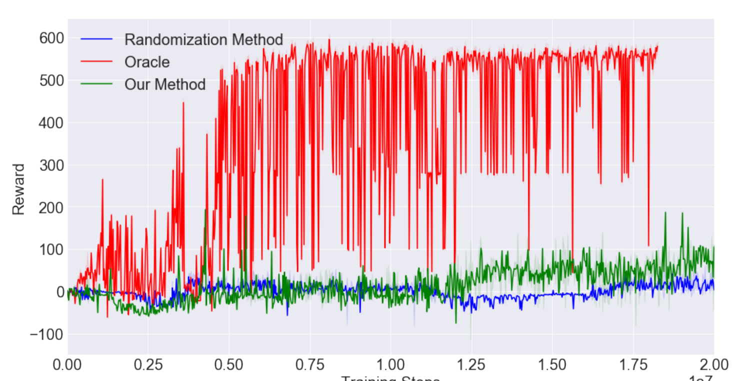
Virtual to Real Reinforcement Learning on Real World Driving Data

We train our RL model with a trained realistic translation network in virtual environment, and tested on a real world driving data to evaluate its steering angle prediction accuracy, results shown as follows.

	Ours	B-RL	SV
Accuracy	43.40%	28.33%	53.60%

Results show that our proposed method outperforms the baseline method (B-RL), where the reinforcement training agent is trained in a virtual environment without seeing any real data. The supervised method (SV) has the best overall performance, however, was trained with large amounts of supervised labeled data.

Transfer learning in Different Virtual Environments



In this experiment, both the baseline (Randomization Method) and our model are trained in the Cg−track2 track and evaluate in E-track1 track, which has different visual appearance.
Obviously, standard A3C (Oracle) trained and tested in the same environment gets the best performance. However, our model performs better than the domain randomization method, which requires training in multiple environments to generalize.

Reference

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. CoRR, abs/1611.07004, 2016. URL http://arxiv.org/abs/1611.07004.
[2] Sully Chen. Autopilot-tensorflow, 2016. URL https://github.com/SullyChen/Autopilot-TensorFlow.
[3] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. arXiv preprint arXiv:1511.00561, 2015.
[4] Bernhard Wymann, Eric Espié, Christophe Guionneau, Christos Dimitrakakis, Rémi Coulom, and Andrew Sumner. Torcs, the open racing car simulator. Software available at http://torcs. sourceforge. net, 2000.

个人信息：游宇榕（计算机科学方向2014级）
邮　　箱：yurongyou@sjtu.edu.cn

致远学院 上海交通大学致远荣誉计划
ZHIYUAN HONORS PROGRAM