

# CycleSquareGAN : 장기간 로봇 자율주행을 위한 이미지 변환

## CycleSquareGAN : Image Translation for Long-term Robot Autonomy

○이 동 현<sup>1</sup>, 김 진 기<sup>1</sup>, 장 우 현<sup>1</sup>, 허 정 범<sup>1</sup>, 오 정 현<sup>1\*</sup>

<sup>1)</sup> 광운대학교 로봇학부 (TEL: 010-3933-9168; E-mail: tjsqlfkdl@snaver.com )

**Abstract** This paper proposes a method to overcome the limitations of existing deep learning models such as pix2pix and CycleGAN, which are used for precise positioning of mobile robots using vision sensors. The proposed method uses the CycleGAN model to transform images from various time periods to a consistent time period, and introduces a Middle Bridge Space (MBS) structure between Generators to improve the independent enhancement of time-dependent images. Additionally, the proposed method uses unpaired image datasets, Discriminator, and SSIM-loss to ensure consistency in recognizing the same place during long-term missions of autonomous driving robots.

**Keywords** Mobile Robot, CycleGAN, Image Translation

### 1. 서론

모바일 로봇이 다양한 서비스 분야에서 활용되면서 비전 센서를 이용한 로봇의 정밀한 위치 추정이 널리 연구되고 있다 [1]. 그러나 실외 환경에서 모바일 로봇의 위치 추정 정확도는 조도의 변화에 영향을 받는다 [2]. 이를 극복하기 위해서 pix2pix[3], CycleGAN[4] 등의 딥러닝 모델을 이용하여 이미지를 변환하는 시도들이 있다. 그러나 [3]은 단일 이미지가 아니라 이미지 pair 가 필요하기 때문에 데이터를 구하기 어렵다는 단점을 가진다. 또한 [4]는 두 개의 도메인만 상호작용할 수 있다는 단점이 존재한다.

본 논문은 기존 모델들의 이러한 한계를 극복하기 위하여 CycleGAN 모델을 통해 다양한 시간대의 이미지를 일정한 시간대의 이미지로 변환하는 방법을 제안한다. 우리는 기존 CycleGAN 의 Generator 의 구조에 변화를 주어 도메인의 변환을 통해 시간에 의존적인 이미지를 독립적으로 개선하였다. 제안한 방법은 Generator 사이에 중간 단계인 Middle Bridge Space 를 구조를 추가하고 unpair 한 이미지 데이터셋을 활용한다. 또한, Discriminator 와 SSIM-loss[5]를 통해 로봇이 장기간 임무 수행 시, 동일한 장소를 일관성 있게 인식할 수 있다. 따라서 자율 주행 로봇의 장기간 임무 수행에 활용될 수 있다.

### 2. 학습 방식

#### 2.1 CycleSquareGAN

본 논문에서 제안하는 CycleSquareGAN 은 CycleGAN[4]을 기반으로 한다. CycleGAN은 두 개의 생성자와(Generator)와 두 개의 판별자(Discriminator)를 사용하여 두 도메인 간의 이미지 변환을 수행할 수 있다. CycleGAN 을 기반으로 하여 unpaired 한 데이터셋을 사용하여 학습할 수 있고, cycle-consistency 를 통해 입력의 특징을 보존할 수 있지만 두 개의 도메인만 상호작용할 수 있다는 단점이 존재한다. 이 문제를 극복하기 위해 CycleSquareGAN 네트워크는 그림 1과 같이 기존 CycleGAN 구조에서 생성자  $G_M$ 을 추가하였다.  $G_M$ 은 여러 시간 대의 이미지들을 공동의 도메인으로 만들어 주기 위한 역할을 하며 이때의 출력을 Middle Bridge Space라 한다.  $G_A$ ,  $G_B$ 는 Middle Bridge Space 를 각각 도메인 A 와 도메인 B 의 이미지로 만들어주는 생성자를 말한다.  $D_A$ ,  $D_B$ 는 각각 입력 이미지를 도메인 A 와 도메인 B 에 속하는지 판별한다. 이때 A, B는 변경대상 시간 도메인과 목표하는 시간 도메인이다.

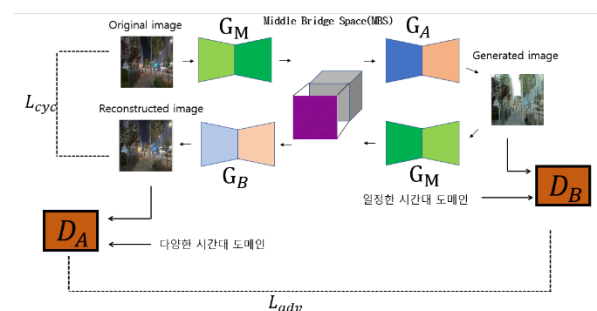


그림 1. 제안하는 학습 모델

\* 본 논문은 교육부와 한국연구재단의 지원으로 지원을 받아 수행된 첨단분야 혁신융합대학사업의 연구결과입니다.

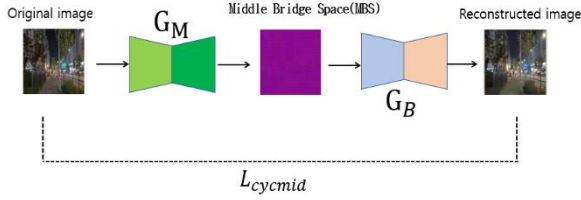


그림 2. 제안하는 학습 모델의 입력 도메인과 Middle Bridge Space 간의 cycle Loss

## 2.2 Formulation

$L_{adv}$ (Adversarial Loss)는 판별자가 생성된 이미지를 진짜 이미지와 구분할 수 없을 정도로 생성자가 이미지를 생성하도록 유도한다. 수식은 (1)과 같다.

$$\begin{aligned} \mathcal{L}_{adv}(G_A, G_B, D_A, D_B, G_M, A, B) = & \mathbb{E}_{A \sim P_{data}(A)} [\log D_A(A)] \\ & + \mathbb{E}_{B \sim P_{data}(B)} [\log D_B(B)] \\ & + \mathbb{E}_{A \sim P_{data}(A)} [\log(1 - D_B(G_A(G_M(A))))] \\ & + \mathbb{E}_{B \sim P_{data}(B)} [\log(1 - D_A(G_B(G_M(B))))] \quad (1) \end{aligned}$$

$L_{cyc}$ (Cycle-Consistency Loss)는 전체 cycle Loss로 각각 생성한 이미지를 다시 원본으로 복구할 때 원본과 복구 값 간의 거리를 구하는 것으로, 생성된 이미지가 다시 원본으로 대응될 수 있게끔 한다. 그림 1에서 도메인 A를 입력으로  $G_M, G_A$ 를 통해 만들어진 가짜 B 데이터를 역순으로  $G_M, G_B$ 를 통해 재구성한 A 이미지와 원본 A 간의 L1 Loss를 계산한다. 도메인 B에 대해서도 동일한 계산을 진행한다. 수식은(2)와 같다.

$$\begin{aligned} \mathcal{L}_{cyc}(G_A, G_B, G_M) = & \|G_B(G_M(G_A(A))) - A\| \\ & + \|G_A(G_M(G_B(B))) - B\| \quad (2) \end{aligned}$$

$L_{cycmid}$ (Cycle-Consistency Loss)는 cycle Loss로, 그림 2와 같이 도메인 A에서 만든 mid를 통해 재구성한 A 데이터와 실제 데이터 A에 대한 L1 Loss[4]를 사용하였다. 이 역시 도메인 B에 대해서도 동일한 계산을 진행한다. 수식은 (3)과 같다.

$$\begin{aligned} \mathcal{L}_{cycmid}(G_A, G_B, G_M) = & \|G_B(G_M(A)) - A\| \\ & + \|G_A(G_M(B)) - B\| \quad (3) \end{aligned}$$

## 3. 실험

### 3.1 GAN model 성능 비교

같은 데이터 셋으로 학습시킨 pix2pix와 cycleGAN, CycleSquareGAN의 각각의 test 결과를 이용해 FID, SSIM 등을 이용해 모델의 성능평가를 비교한다.

표 1. 모델 비교

	Pix2Pix	CycleGAN	CycleSquareGAN
InceptionV3	1.034	1.044	<b>1.112</b>
FID	260.984	<b>92.616</b>	95.900
SSIM	0.002	0.004	<b>0.005</b>



그림 3. Test 결과

제안된 모델의 경우 CycleGAN구조를 기반으로 새로운 모델을 구축하였기 때문에 CycleGAN구조와 성능 비교에서 비슷한 성능을 보여줄 수 있다.

## 4. 결론

본 논문에서는 제안하는 CycleSquareGAN 모델은 CycleGAN의 구조에 MBS와 SSIM-loss를 추가한 모델로 시간에 의존적이라는 한계를 극복하였다. 또한, 앞서 제기되었던 모델들 보다 유사하거나 더 높은 성능을 보인다.

## 참고문헌

- [1] Scaramuzza, D., Martinelli, A., & Siegwart, R. (2006). Visual Teach and Repeat for Mobile Robots. In Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA) (pp. 118-123).
- [2] Lee, J. H., Lim, H., Park, C., & Kim, H. J. Effects of Illuminance Changes on Mobile Robot Localization Accuracy in Outdoor Environments. International Journal of Distributed Sensor Networks, 11(3), 925854.
- [3] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3-15).
- [4] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2223-2232).
- [5] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 13(4), 600-612.