

불균형 학습을 개선한 Cycle-GAN과 부분적 Stable Diffusion을 활용한 고 품질의 가상 악천후 도로 이미지 생성

윤정현·김주희*

국립창원대학교 메카트로닉스공학부

Generating Virtual Adverse Weather Images with High-Quality Using an Integrated Learning-Balanced Cycle-GAN and Partial Stable Diffusion

Jeonghyeon Yoon · Juhui Gim*

School of Mechatronics Engineering, Changwon National University, Gyeongnam 51140, Korea

(Received 22 November 2024 / Revised 23 December 2024 / Accepted 23 December 2024)

Abstract : This paper proposes a virtual yet high quality image of adverse weather conditions that minimizes information loss and distortion. The proposed method comprises two main components: the development of a Cycle-GAN technique that preserves image quality by reducing the performance of discriminators and by modifying the network structure of generators, and a partial stable diffusion technique that transforms only the targeted areas, utilizing Mask R-CNN for precise region identification. High-quality image generation under various adverse weather conditions can be achieved by transferring features through the proposed Cycle-GAN and by changing the structural features of the road in the image with partial stable diffusion. Compared to the original Cycle-GAN and stable diffusion, the proposed algorithm generates a variety of adverse weather road images without compromising quality. By applying the augmented database created through this algorithm to a deep learning-based road area detection model, the results demonstrate improved performance in deep learning-based detection tasks.

Key words : Virtual image generation(가상 이미지 생성), Learning-balanced Cycle-GAN(학습 균형 사이클 간), Partial stable diffusion(부분적 안정 확산 기법), Data augmentation(데이터 증강), Adverse weather conditions(악천후 환경)

1. 서론

정확한 주변 환경 인지는 차량을 완벽히 제어하는데 필수적인 기술이다.^{1,4)} 딥러닝 기술과 컴퓨터 성능의 발전으로 차량의 주변 환경 인식 성능은 지속적으로 향상하고 있다. 하지만 딥러닝 기반 인식 기술은 여전히 악천후 환경에서는 성능 저하를 겪는다. 그 이유는 딥러닝 기술은 데이터 의존적이며, 악천후 환경에서의 도로에 대한 데이터가 맑은 날 도로에 대한 데이터보다 현저히 부족하기 때문이다.

악천후 환경에서의 딥러닝 기반 인식 성능을 확보하기 위한 가장 근본적인 방법은 다양한 악천후 환경에서의 상황을 골고루 반영한 양질의 데이터를 활용하여 모

델을 학습하는 것이다. 하지만 악천후 환경의 날씨 분포가 불균등하며, 그 날씨에 해당하는 강수량 역시 일정하지 않기 때문에, 완벽한 양질의 데이터를 직접 구축하는 것은 현실적으로 불가능하다. 이로 인해, 악천후 환경에서의 딥러닝 인식 모델의 정확성 및 신뢰성을 보장하는 것은 항상 대두되는 문제다.^{5,6)}

데이터의 직접 구축 방법에 대한 한계를 해결하기 위한 방법으로 이미지 확대 및 축소, 색상과 밝기 조절 등을 통해 원본 데이터를 변형하는 데이터 증강 방식이 있다.⁷⁾ 하지만 이 방식은 원본 데이터의 근본적인 특징을 변경하지 못하기 때문에 다양한 날씨를 구현하지 못한다는 단점이 있다. 따라서 악천후 환경에 대한 다양성,

*Corresponding author, E-mail: juhuigim@changwon.ac.kr

[†]This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

그에 따른 학습 성능을 증가시키는 데는 한계가 있다.

최근에는 생성형 AI(Generative AI) 기법이 활발히 연구되면서, 데이터 증강을 위해서도 딥러닝 기법이 활용되고 있다.⁸⁻¹⁸⁾ 생성형 AI 기법은 원본 데이터의 특징을 변형하거나 특징 간 조합을 통해 학습 데이터의 다양성과 양을 인위적으로 증폭시킨다. 따라서 생성형 AI 기법을 통한 가상 이미지 데이터 생성은 데이터 확보의 한계를 극복하는 효과적인 해결책으로 대두된다. 데이터 생성 시 가장 많이 활용되는 생성형 AI 기법은 Generative Adversarial Nets(GAN) 기반 방법과 Diffusion model 기반 방법이 있다.^{19,20)}

GAN을 활용한 가상 이미지 생성 방식은 특징 전달에 강화된 Cycle-GAN 및 Star-GAN 등이 주로 활용되어왔다.⁸⁻¹²⁾ 그러나 Cycle-GAN과 Star-GAN 기법은 이미지 구조를 유지하면서 이미지 특징을 전달하는데 그 목적이 있기 때문에, 다양한 이미지 특징 변환(Translation)은 가능하나 다양한 이미지 구조 생성(Transformation)에는 한계가 있다.²¹⁻²⁵⁾ 또한, 적대적 학습을 통해 가상 이미지를 생성하는 GAN 기법의 특징 상, 생성자(Generator)와 판별자(Discriminator)의 학습 균형이 가상 이미지 데이터 품질에 중요한 영향을 끼치는데, 두 경쟁 네트워크의 학습 균형을 유지하는 것은 매우 어렵다.²⁵⁻³¹⁾

Diffusion model을 활용한 가상 이미지 생성 방식에서는 텍스트 프롬프트를 통해 요구되는 특징을 안정하게 생성 가능한 Stable diffusion이 활용된다.³²⁾ 따라서 이미지 구조 생성에는 Stable diffusion 방식이 Cycle 및 Star-GAN 기반 방식 대비 좀 더 적합하다. 하지만 텍스트 프롬프트 기반 특징 입력 방식은 텍스트와 이미지 간의 연계가 쉽지 않아 세부적인 이미지 변경이 어렵고, 입력 받은 특징을 충족하기 위하여 오히려 원본 이미지의 필수적인 정보마저 손실되는 현상이 발생한다.^{33,34)}

생성형 AI 기법을 활용한 양질의 데이터를 생성하기 위해서는 악천후 환경 특징을 원하는 영역에 잘 반영할 수 있도록 해야 한다. 하지만 이미지 특징 전달에 강화된 GAN 기반 기법의 경우, 원본 이미지에 악천후로 인한 환경 변화 특징 전달은 잘 되지만, 눈 쌓임 정도 및 젖음 정도 등 특징 전달 수준을 조절하지 못해 악천후 환경에 대한 다양성을 확보하는 것은 불가능하다. Stable diffusion 기법의 경우 프롬프트를 통해 악천후 환경에 대한 다양성을 확보할 수는 있지만, 건물 및 주변 도로 정보 등 악천후 이외의 다른 도로의 특징까지 생성 및 변형해버린다는 단점이 있다. 그러므로 악천후 환경에 대한 다양성을 확보하면서도 중요 도로 특징은 유지하는 데이터 생성 기법 개발이 요구된다.

본 연구는 이미지 품질 손상 없이 악천후 환경에 대한

특징만을 전달하면서도 악천후 환경의 다양성을 생성할 수 있는 도로 이미지 생성 알고리즘을 제안한다. 제안하는 알고리즘은 경쟁 네트워크 간 학습 균형 문제를 해결한 Cycle-GAN으로 악천후 환경에 대한 분위기를 포함하는 특징을 전달하고, Mask R-CNN을 통해 추출된 도로 영역을 대상으로 Stable diffusion을 적용함으로써 악천후 환경에 대한 도로 환경 조건을 변경하여 가상 이미지를 최종적으로 생성한다. 제안된 방법론은 기존 가상 이미지 생성에서의 단점을 두 개의 측면에서 해결하여, 품질을 유지하면서도 정보 손실을 최소화하는 가상 이미지 생성을 가능하게 한다는 것을 보여준다. 첫째, 이미지 특징 전달 시 이미지 품질을 저하시키는 Cycle-GAN의 생성자와 판별자의 학습 균형 문제를 해결하여 열악한 학습 조건에서도 이미지 품질을 유지한다. 둘째, 엄격하게 제어되지 않은 텍스트 프롬프트에 의해 변경하고자 하는 영역 이외의 필수적인 정보 보존이 어려운 Stable diffusion의 문제를 해결하기 위해 변경하고자 하는 영역을 감지하여, 일부 영역에만 Stable diffusion을 적용하는 Patial stable diffusion을 사용함으로써 필수적인 정보를 보존한다. 제안된 알고리즘을 통해 악천후에 대해 생성한 이미지를 활용하여 도로 노면 추출 알고리즘에 적용한 결과, 악천후 도로 상황에서 우수한 도로 노면 추출 결과를 제시하여 본 알고리즘의 적용 가능성을 입증한다.

본 논문의 구조는 다음과 같다. 2장에서는 본 알고리즘을 구성하는 주요 딥러닝 기법들을 설명하고, 3장에서는 제안하는 알고리즘을 설명한다. 4장에서는 제안하는 알고리즘의 성능을 검증함과 동시에, 생성된 가상 이미지 기반의 데이터베이스를 활용하여 악천후 환경에서의 도로 노면을 검출함으로써 딥러닝의 인지 성능을 강화할 수 있음을 보여준다. 그리고 5장 결론으로 마무리한다.

2. 관련 연구

이 장에서는 본 알고리즘 개발 시 활용된 Cycle-GAN, Mask R-CNN, 그리고 Stable diffusion의 개념을 설명한다.

2.1 Cycle-GAN

Cycle-GAN은 가상 이미지를 생성하는 생성자(Generator)와 생성된 이미지의 진위 여부를 판별하는 판별자(Discriminator) 간 적대적 학습을 통해 가상의 이미지를 생성하는 GAN 구조를 기반으로 이미지 특징을 전달하는 Unpaired image-to-image translation 기술 중 하나이다.³⁵⁾ Fig. 1은 Cycle-GAN의 원 구조를 나타내며, 두 원본 이미지 X, Y 각각의 구조적 정보 일관성을 유지하면서도 서로 간의 특징만이 전달된 Y', X' 를 생성한다.

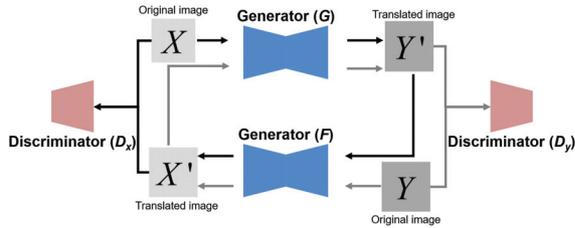


Fig. 1 Framework of Cycle-GAN

2.2 Mask R-CNN

Mask R-CNN은 이미지 내 개별 객체의 경계를 마스크 형태로 구분하는 Instance segmentation 기반 기법 중 하나이다.³⁶⁾ Fig. 2는 Mask R-CNN의 기본 구조를 나타낸다. 입력 이미지는 세부적인 분석을 수행하는 Top-down 방식의 네트워크와 물체를 픽셀 단위 특징을 영역 단위로 확장하는 Bottom-up 방식의 네트워크가 통합된 Pyramid 구조를 통해 특징 맵(Feature map)으로 추출된다. 특징 맵을 구성하는 객체 후보 영역(Region proposal)들 중 Region Proposal Network(RPN)에 의해 선정된 최종 객체 후보 영역의 정보를 RoI align을 통해 세밀화 한다. 최종적으로 세밀화된 객체 후보 영역에 대해 객체 분류(Classification), 객체 위치를 나타내는 바운딩 박스 생성(Bounding box), 그리고 객체 영역으로의 마스크 예측(Mask prediction)을 병렬로 처리하여, 출력된 객체 후보를 마스크로 출력한다.

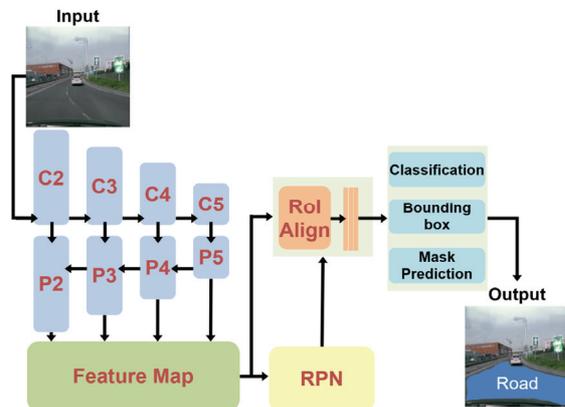


Fig. 2 Framework of Mask R-CNN

2.3 Stable Diffusion

Stable diffusion은 Diffusion model의 일종으로, 텍스트와 이미지 간의 연관성을 추출하여 프롬프트로 입력된 특징을 통해 이미지를 변경(Transformation)한다.³²⁾ Fig. 3은 Stable diffusion의 기본적인 구조를 보여준다. 입력 이

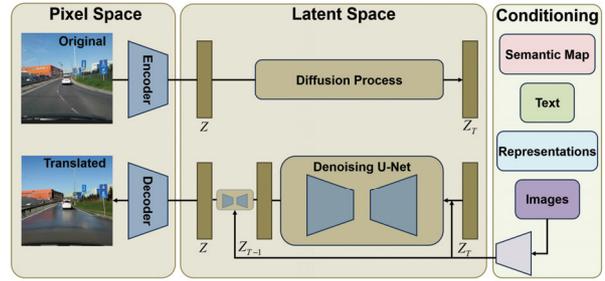


Fig. 3 Framework of Stable diffusion

미지는 Variational autoencoder(VAE) 모델의 인코더(Encoder)를 통해 잠재 공간(Latent space)으로 압축된 후, 잠재 벡터(Z)로 변환된다. 고 수준의 특징을 포함한 잠재 벡터(Z)는 Diffusion 과정을 거쳐 특징 변경이 용이한 완전한 노이즈 형태의 잠재 벡터(Z_T)로 변형된다. 변형된 잠재 벡터(Z_T)는 Semantic map, Text prompt, Representations, Images로부터 입력 받은 원하는 특징을 포함하면서 점차 노이즈가 제거된 잠재 벡터(Z)로 전환되고, 최종적으로 VAE의 디코더(Decoder)를 통과하면서 원하는 특징을 반영한 이미지로 복원된다.

3. 방법론

제안하는 품질 손상 없이 악천후에 대한 다양한 도로 환경 이미지를 생성하는 알고리즘은 Fig. 4와 같이 악천후 환경의 특징적 변경 및 부분 구조적 변경의 직렬 연결 구조로 구성된다. 특징적 변경 과정에서는 생성자-판별자 간 불균형 학습 문제를 완화한 Cycle-GAN을 활용하여, 이미지 품질 손상 없이 악천후 환경으로의 전반적인 분위기 변화를 수행한다. 구조적 변경 과정은 요구되는 영역 이외의 구조적 정보도 변경되는 Stable diffusion 한계를 해결하기 위해 Partial stable diffusion을 활용하여 악천후 환경에 해당하는 도로 노면 특징 변경을 수행한다.

3.1 Feature Translation with Learning-balanced Cycle-GAN

Cycle-GAN기법 모델은 맑은 날의 도로 이미지(X)의 전체적인 분위기를 악천후 환경(Y)의 분위기로 변환한다. 하지만 생성자와 판별자의 적대적인 학습을 하는 GAN 기법의 특징 상, 생성자-판별자 간 구조적 복잡성 차이로 인해 생성한 이미지의 품질이 급격히 저하되는 단점이 있다.^{37,38)}

생성자는 이미지의 진위 구조만 판별하여 단순한 컨볼루션 구조를 지닌 판별자와 달리, 이미지를 생성해야 하므로 인코더-디코더로 구성된 복잡한 오토 인코더 구조를 지니고 있기 때문이다. 따라서 생성자는 기본적으로

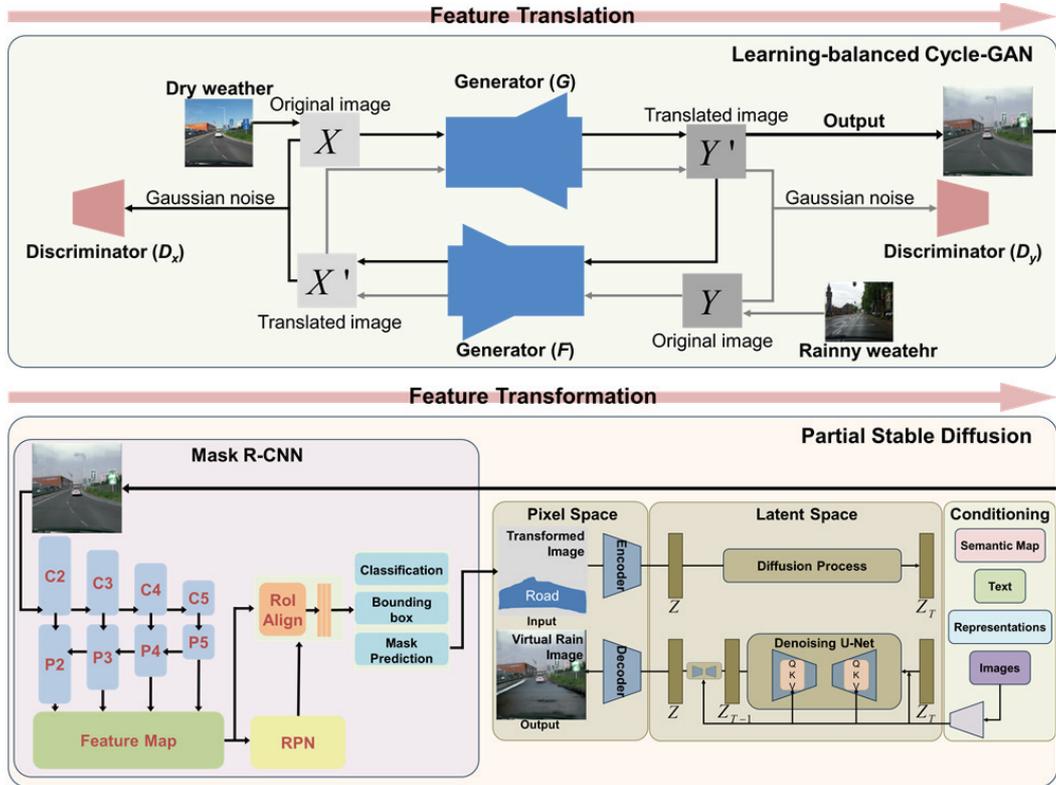


Fig. 4 Overall structure of the proposed virtual image generator

로 판별자보다 파라미터 용량이 크고 학습 속도가 느려, 생성자가 판별자의 피드백을 올바르게 반영하지 못한다. 특히, 이 문제는 악천후에서의 이미지 데이터베이스와 같이 데이터 불균형 문제가 심할수록 두드러진다. 본 연구에서는 실제 원본의 이미지와 유사한 품질을 지닌 가상 이미지 생성을 위해 기존 Cycle-GAN의 단점들 1) 판별자 성능 저하와 2) Cycle-GAN의 생성자 네트워크 구조 변경을 통해 해결한다.

3.1.1 Discriminator Performance Reduction

판별자의 성능 저하를 위한 첫 번째 방법은 판별자의 학습률(Learning rate)을 조절하는 것이다. 일반적으로, Cycle-GAN의 판별자와 생성자의 학습률은 비슷한 속도로 설정된다. 비슷한 속도로 설정된 학습률로 인해, 생성자 대비 상대적으로 파라미터 용량이 적은 판별자는 Fig. 5(a)와 같이 빠른 속도로 수렴되어, 생성자가 제대로 된 이미지를 생성하기도 전에, 판별자가 생성 결과물의 진위 여부를 명확하게 판단한다.

Fig. 5(b)는 판별자의 학습률을 생성자의 학습률보다 높게 설정했을 때의 손실 값을 보여준다. 초기에는 제안된 학습률에 의해 판별자의 초기 학습이 불안정해져 손실이 초기값 대비 상승하고, 이로 인해 수렴 속도가 Fig.

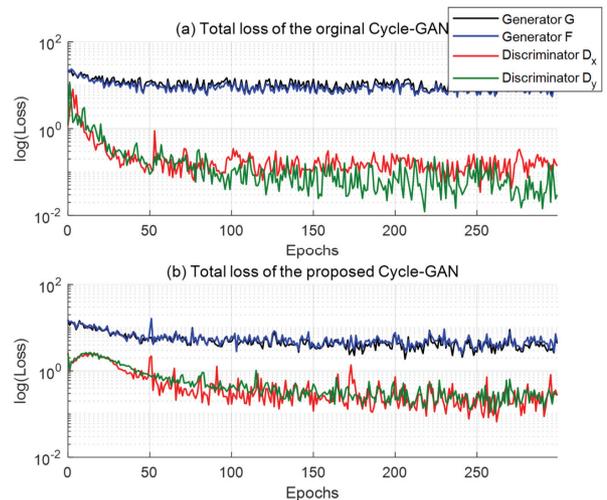


Fig. 5 Loss of the generators and discriminators: (a) in the original Cycle-GAN, and (b) in the proposed Cycle-GAN

5(a) 대비 느려짐을 확인할 수 있다. 생성자가 수렴되기 전에 판별자의 수렴 속도를 늦춰 손실 값의 차이를 줄임으로써, 생성자가 입력 이미지의 품질을 유지할 수 있도록 한다.

판별자의 성능 저하를 위한 두 번째 방법은 Fig. 6과

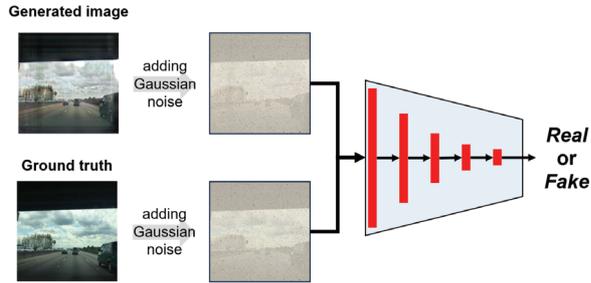


Fig. 6 Discriminator training with the Gaussian noise-added images

같이 판별자의 훈련을 위한 입력 이미지에 노이즈를 삽입하여 훈련하는 것이다. 노이즈가 포함된 이미지를 사용하여 판별자를 훈련하는 방법은 생성자가 낮은 품질의 이미지를 생성하였을 때 판별자가 쉽게 판단하지 못하게 하는 방법으로, 판별자의 판단 능력을 저하시켜, 향상된 피드백 과정을 통해 생성자의 가상 이미지 생성 능력을 높인다.

3.1.2 Generator Network Design

생성된 이미지의 품질 저하 문제를 해결하기 위해, 생성자 구조 설계의 기본적인 개념은 Fig. 7과 같이 이미지 품질을 저하시킬 수 있는 네트워크를 최대한 제거하는 것이다. Down-sampling을 통해 이미지의 특징을 추출하는 인코더는 특징을 집약하면서도 동시에 압축 과정에서 원본 해상도 손실되는 문제가 발생한다. 제안하는 생성자 구조는 인코더에서 특징 맵을 압축시키는 Down-sampling을 제거하여 원본과 같은 크기의 특징 맵을 추출하고, 이미지로 복원하는 디코더에서 기존 이미지 크기 대비 더 큰 크기로 Up-sampling을 진행하여 특징 맵의 해상도를 향상시킨다. 이후, Resizing 과정을 통해 향상된 해상도를 원본의 이미지 크기로 다시 조정함으로써 이미지 품질을 유지하도록 한다.

Up-sampling 디코더와 Resizing의 조합은 생성자의 이미지 생성 과정에서 세부 정보를 효과적으로 재구성하게 한다. 또한, Fig. 5(b)의 생성자 손실 값에서 확인할 수 있듯이, 기존 Cycle-GAN 대비 생성자의 학습 손실을 안정적으로 유지하는 데 기여한다.

3.2 Feature Transformation with Partial Stable Diffusion

Stable diffusion을 활용한 구조적 변경의 목적은, 악천후 환경의 특징적 변경이 완료된 이미지를 대상으로 도로 노면 위 쌓인 눈 높낮이 조절 및 비에 의한 노면 반사 등을 구현하는 것이다. 그러나 기존의 Stable diffusion 기

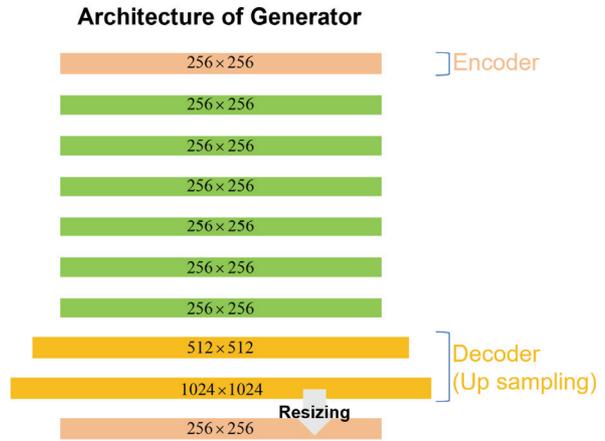


Fig. 7 Proposed structure of generators in the proposed Cycle-GAN

법을 적용하면, 도로 환경에서 날씨 영향을 받지 않는 물체들까지 구조를 변경시킨다는 단점이 발생한다. 원하는 영역에 대해서만 악천후 환경에 대한 구조적 변경을 수행하기 위하여 1) Mask R-CNN을 통한 이미지 변경 영역 추출과 2) 추출된 이미지 변경 영역에 대한 Stable diffusion 적용을 진행한다.

3.2.1 Image Transform Area Extraction

이미지 변경 영역 추출을 위한 모델은 Mask R-CNN ResNet50-FPN을 활용한다. 맑은 날씨에 대한 도로 환경 이미지 데이터를 통해 도로 노면 정보를 추출하도록 학습된 Mask R-CNN은 RPN을 활용하여 입력 받은 도로 이미지에서 도로 노면을 분류 및 위치를 추정하고 추정된 도로 노면을 관심 영역(Region of interests)으로 지정한다. 최종적으로 추출된 도로 노면 결과는 Stable diffusion이 적용될 구조 변경 영역으로 활용된다.

3.2.2 Partial Stable Diffusion

Mask R-CNN을 통해 추출된 도로 영역에 대해서만 Stable diffusion을 적용하여 악천후 환경에서의 도로 상태로 특징 구조를 변경한다. 기존의 Stable diffusion을 활용하여 일부 영역만 변경하려면, 변경 및 유지하고자 하는 구조에 대한 자세하고 긴 텍스트 프롬프트가 요구된다. 그러나 긴 텍스트 프롬프트는 Stable diffusion이 내용을 제대로 이해하지 못할 수도 있으며, 오역으로 인해 변경을 하고자 하는 영역 이외의 정보가 변형되어 정보 손실 발생하는 가능성이 높다. 반면, 제한된 영역 기반 Stable diffusion은 변경을 원하지 않은 영역에 대한 지시 사항을 무시할 수 있어, 프롬프트를 더욱 직관적이고 간

략하게 제시 가능하다.

단순해진 악천후 환경에 대한 도로 상태 특징 구조 변경을 위한 프롬프트는 두 가지 조건의 조합으로 구성된다. 첫번째 조건은 ‘Asphalt’, ‘Road’, ‘Off-road’와 같은 도로 노면 상태를 나타내는 조건이며, 두번째 조건은 ‘Wet’, ‘Snowy’ 등과 같은 날씨 및 대기 상태를 표현하는 조건이다. 단순화된 프롬프트로 인해, 제안하는 부분적 Stable diffusion 적용 결과는 기존 결과와 달리 이미지 변환 강도에 대한 민감도가 낮으며, 이미지 변환 강도 정도를 조절함으로써 눈 쌓임 정도, 비가 내리는 정도 등 다양한 날씨 환경 구성을 가능하게 한다.

4. 검증 및 적용

4.1 Implementation

본 연구에서 사용한 학습 환경은 RTX4090 그래픽을 기반의 워크스테이션을 이용하였다. 또한, 이미지 특징 전달을 위한 Cycle-GAN 학습에는 256 × 256 크기의 맑은 날 이미지 300장, 우천시 이미지 72장, 밤 이미지 42장을 학습하였다. 또한, 도로 이미지 관심영역 추출을 위한 Mask R-CNN 모델 학습을 위한 데이터로는 256 × 256 크기의 맑은 날 도로 이미지 253장과 우천시 도로 이미지 12장 및 밤 이미지 19장을 사용하여 학습을 진행하였다.

4.2 Validation

Fig. 8은 두 개의 맑은 날의 도로 환경 이미지(Original)를 기존 Cycle-GAN과 제안한 방법의 Cycle-GAN(Proposed)을 활용하여 비가 오는 날의 이미지로 변환한 결과를 보여준다. 기존 Cycle-GAN이 생성한 이미지의 경우, 가장자리 영역 및 차량과 나무와 같은 물체 일부에서 이미지의 픽셀이 뭉치는 아티팩트(Artifact) 현상과 생성된 이미지의 물체가 모드 붕괴(Mode collapse) 현상이 발생한 것을 보인다. 반면, 제안된 Cycle-GAN의 결과는 아티팩트 및 모드 붕괴현상이 상대적으로 적은, 기존 Cycle-GAN 대비 높은 품질의 가상 이미지를 생성한다. 따라서, 제안된 Cycle-GAN이 기존 Cycle-GAN 대비 품질 손상 없이 이미지 특징을 훨씬 잘 전달하는 것을 확인할 수 있다.

Fig. 9는 Partial stable diffusion의 이미지 변경 영역 지정 구간에서 활용한 Mask R-CNN의 감지 성능을 보여준다. 정성적인 평가와 Intersection over Union(IoU) 기반의 정량적인 평가를 통해 맑은 날 데이터로 학습한 Mask R-CNN 모델이 변경할 구역을 높은 정확도로 감지하는 것을 검증하였다. 최종적으로 Mask R-CNN으로 추출된 영역을 대상으로 Stable diffusion이 적용된다.

Figs. 10과 11은 각각 맑은 날 도로 이미지를 우천시와

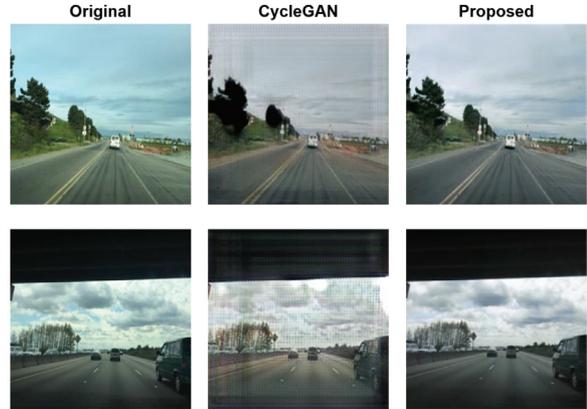


Fig. 8 Performance of the proposed Cycle-GAN

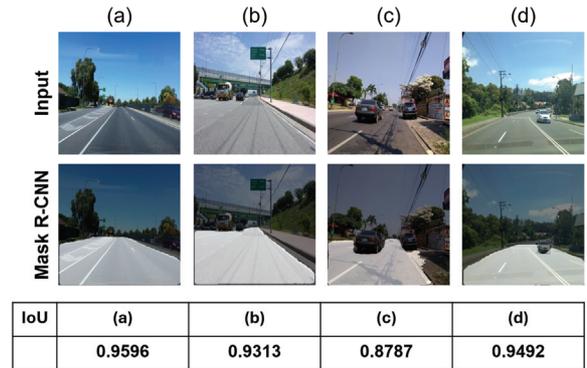


Fig. 9 Performance of the mask R-CNN in partial stable diffusion

설한지 도로 이미지로 변환한 결과를 0.4, 0.7의 변환 강도를 가진 Stable diffusion, 제안된 Cycle-GAN, 그리고 제안된 Cycle-GAN 결과에 부분 Stable diffusion을 적용하는 제안하는 알고리즘으로 변환한 이미지를 보여준다. Stable diffusion의 텍스트 프롬프트는 ‘Wet asphalt, Cloudy sky’, ‘Snowy asphalt, Cloudy sky’와 같이 단순하게 설정하였다. 제약 조건이 단순하게 설정된 텍스트 프롬프트를 기반으로 생성한 이미지들은 Stable diffusion에 의해 도로와 하늘 영역 이외에 정보도 함께 왜곡되고, 교통 신호등이나 건물과 같은 구조물이 생성되거나 소실된 것을 확인할 수 있다. 예를 들어, Fig. 11의 세 번째 이미지 결과에서 Stable diffusion을 통해 변경한 이미지들의 경우, 변환 강도 0.4의 경우 주변 집이 다른 형태의 집으로 변경되었으며, 변환 강도 0.7의 경우 건물이 나무와 같은 숲으로 왜곡된 현상을 확인할 수 있다. 또한, Cycle-GAN의 경우 전체적인 분위기는 잘 변경되었으나, 도로에 눈이 쌓이는 구조적 정보는 거의 형성되지 않은

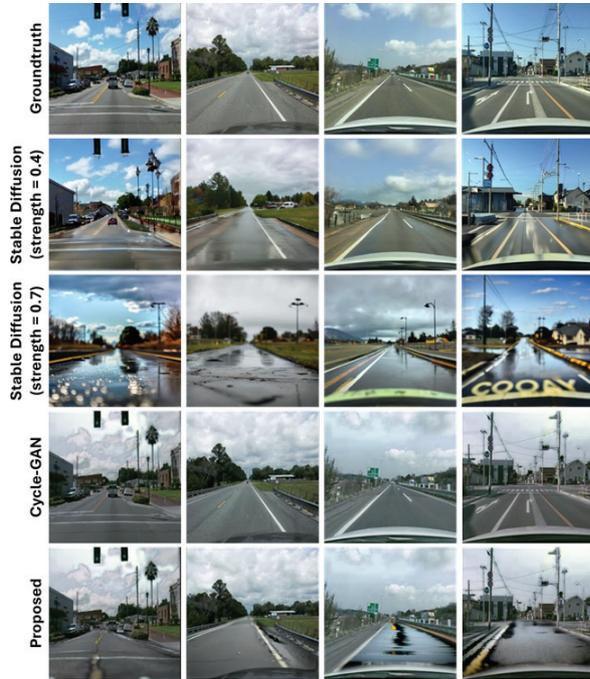


Fig. 10 Generation of virtual road images in rainy days

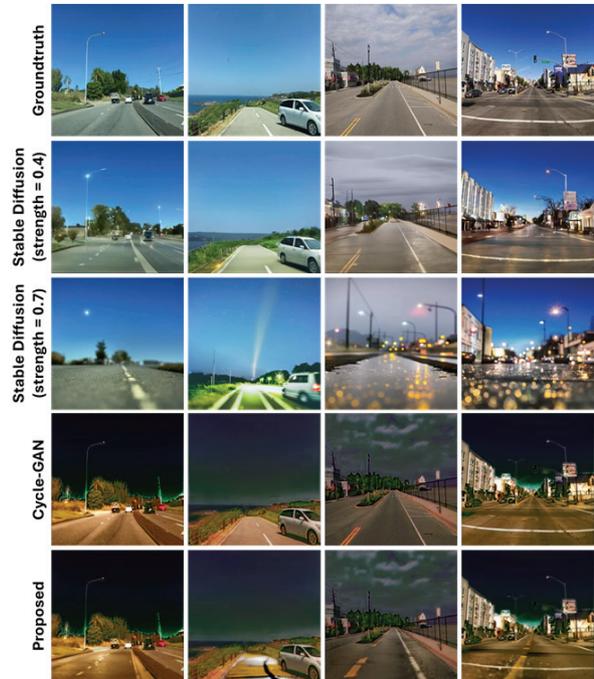


Fig. 12 Generation of virtual night road images in rainy days



Fig. 11 Generation of virtual road images in snowy days

것을 확인할 수 있다. 반면, 제안된 알고리즘의 경우 도로 영역 이외의 건물이나 전봇대가 유지되고 있으며, 도로 위 눈의 쌓임이 잘 표현된 것을 확인할 수 있다. 이 현

상은 변환 강도가 높을수록 더 빈번하게 발생한다.

제안된 Cycle-GAN으로 생성된 이미지들은 맑은 날 이미지에 대해 이미지 품질 손상 없이 비나 눈의 전체적인 분위기를 전달한 채 생성된 것을 확인할 수 있다. 하지만 Cycle-GAN의 특성상 구조적 변경이 제한적이기 때문에, 도로 위 비나 눈의 특징이 잘 반영되지 않은 것을 확인할 수 있다.

제안한 알고리즘(Proposed)의 결과는 제안된 Cycle-GAN의 생성된 이미지를 기반으로 변경 영역인 도로 노면에 대해서 비나 눈의 특징을 ‘Wet asphalt’와 ‘Snowy asphalt’와 같이 단순하게 설정된 프롬프트를 사용하여 생성한 이미지이다. 제안한 방법의 결과는 Stable diffusion 결과 대비 도로 영역 이외의 정보는 유지한 채 비나 눈의 특징을 반영한 것을 확인할 수 있고, 제안한 Cycle-GAN 결과 대비 도로 영역에 대해 비나 눈의 구조적 특징을 잘 반영한 것을 확인할 수 있다.

Fig. 12는 맑은 날 도로 이미지(Ground truth)를 야간의 우천 환경에서의 도로 이미지로 변경한 결과이다. Stable diffusion의 경우 ‘Wet asphalt, night sky’ 프롬프트가 적용되었다. 맑은 낮 이미지에서 흐린 낮 이미지로 변경하는 경우와 달리, 맑은 낮 이미지에서 흐린 밤 이미지로 변경하는 경우, 일정한 가우시안 노이즈 분포를 활용하는 Stable diffusion의 특징 상 밤을 나타내는 어두운 표현이 잘 생성되지 않는다. 반면, 제안하는 알고리즘은 제안하

는 Cycle-GAN 구조를 이용하여 이미지 손상 없이 밤을 나타내는 어두운 특징 전달 후 부분 영역에서 Stable diffusion이 적용되므로, 전체 영역에서 Stable diffusion을 적용한 결과보다 훨씬 밤에 대한 분위기를 잘 표현하면서 비에 의해 젖은 도로 환경을 잘 구현해낸다.

4.3 Application: Deep Learning-based Road Area Detection in Adverse Weather Conditions

제안한 알고리즘을 통해 다양한 악천후 환경에 대하여 고품질의 가상 도로 이미지 데이터를 생성할 수 있다. 본 연구에서는 제안하는 알고리즘을 활용해 증강된 데이터베이스를 도로 영역 탐지에 활용하여, 제안된 알고리즘이 악천후 환경에서의 딥러닝 기반 인지 성능 향상에 기여할 수 있음을 검증한다. 이때 도로 영역 탐지를 위한 딥러닝 모델은 Mask R-CNN ResNet50-FPN을 사용하였다.

원본 데이터베이스의 경우 맑은 날 낮 이미지 253장, 우천 시 낮 이미지 12장, 우천 시 밤 이미지 31장으로 불균형 하게 구성되어 있는 것을 활용하였다. 그리고 제안된 알고리즘을 활용하여 총 2개의 증강 데이터베이스를 구축하였다. 첫 번째 증강 데이터베이스의 경우, 낮 환경에 대하여 맑은 날 비 오는 날에 대한 데이터 균형을 맞춘 데이터베이스로, 원본에서의 낮 이미지(맑은 날 낮 이미지 253장, 우천 시 낮 이미지 12장)와 가상의 우천 시 낮 이미지 241장으로 구성되어 있다. 두 번째 증강 데이터베이스의 경우, 맑은 날 낮 이미지와 비 오는 날 밤 이미지 간 데이터 균형을 맞춘 데이터베이스로, 맑은 날 낮 이미지 253장, 우천 시 밤 이미지 31장과 함께 가상의 우천 시 밤 이미지 222장으로 구성되어 있다.

Fig. 13은 학습 데이터에 포함되지 않은 우천 시 낮과 밤의 도로 환경 이미지들을 원본 데이터베이스로 학습한 모델과 제안된 알고리즘을 통해 증강된 데이터베이스 1과 2로 학습한 모델에 넣었을 때 출력된 결과를 보여준다. 균형 잡히지 않은 데이터를 통해 학습한 모델은 비로 인해 도로 영역이 잘 인식되지 않는 상황과 어두운 밤과 비에 젖어 도로가 잘 보이지 않는 두 상황에서 모두 열악한 인식 결과를 보인다. 반면, 증강된 데이터베이스 1은 낮 이미지에 대해서 우천 환경임에도 불구하고 도로 영역을 상대적으로 잘 감지하는 것을 확인할 수 있다. 또한, 증강된 데이터베이스 2는 비 오는 날 밤 이미지에 대해서도 도로 영역을 상대적으로 잘 감지한다. Table 1은 Fig. 13의 결과에 대하여 Intersection over Union(IoU) 지표를 나열한 것이다. 전체적으로 데이터베이스 1 및 2를 기반으로 학습한 도로 노면 추출 모델이 불균형한 데이터베이스를 기반으로 학습한 모델 대비 훨씬 더 높은 값

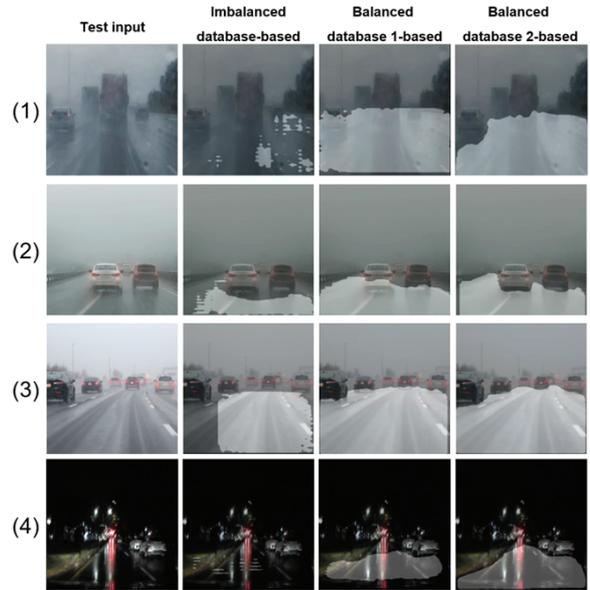


Fig. 13 Application: drivable area detection with the Mask-RCNN trained using the augmented database with the proposed virtual image generator: test input images; results using the original database; results using the augmented daytime database; results using the augmented nighttime database

Table 1 Intersection over Union (IoU) evaluation

Image	Imbalanced dataset-based	Balanced database 1-based	Balanced database 1-based
(1)	0.0775	0.7672	0.8792
(2)	0.5950	0.8147	0.8518
(3)	0.7002	0.9251	0.9587
(4)	0.0986	0.6163	0.6851

을 가지는 것을 정량적으로 확인할 수 있다. 이를 통해 데이터 불균형 문제를 해소하면 딥러닝 기반 인지 성능이 향상된다는 것을 직접 확인하였으며, 또한 제안된 알고리즘이 데이터베이스로 활용 가능함을 검증하였다.

5. 결론

본 연구는 품질 및 필수 정보 손실을 방지한 가상 악천후 환경 이미지 생성 알고리즘을 제안한다. 이를 위해 우선 Cycle-GAN의 판별자 성능 저하 및 생성자 네트워크 구조 변경을 통해 품질 손상을 최소화하면서 원본 이미지에 악천후 환경에 대한 특징을 전달한다. 그리고 Mask R-CNN으로 추출된 도로 노면을 대상으로 부분적 Stable diffusion을 적용함으로써, 간결하고 직관적인 프롬프트

만으로 필수 정보 손실을 최소화하면서도 악천후 환경으로의 도로 환경 구조를 변경한다. 검증 결과, 제안된 방법론은 기존 Cycle-GAN 및 Stable diffusion 대비 높은 품질을 유지하면서도 다양한 조건에서의 악천후 도로 이미지를 생성함을 보였다. 이를 바탕으로 증강된 데이터베이스를 활용하여 도로 영역 탐지에 적용한 결과, 인지 성능이 향상된 것도 확인하였다. 그러나 제안한 방법론 또한 Stable diffusion의 근본적인 한계를 포함하기 때문에, 변경하고자 하는 도로 영역에 대해서 차선의 색이나 종류와 같은 도로 노면의 주요 특징이 변경될 수 있다는 한계점이 존재한다. 그럼에도 불구하고, 제안된 알고리즘은 데이터 획득이 제한된 상황에서도 양질의 데이터 형성을 가능하게 하여, 딥러닝 성능 안정성을 보장한다는 것을 검증하였다. 향후 최적의 텍스트 프롬프트 구성 연구를 수행하여 Partial stable diffusion의 완성도를 향상시키고, 본 알고리즘을 활용하여 생성된 가상 악천후 데이터베이스를 통해, 악천후 상황에서 운전 가능 영역 뿐만 아니라 물체 탐지 및 차선 감지의 성능을 개선하고자 한다.

후 기

이 논문은 2024년도 국립중앙대학교 강사-대학원생-교원 간 융합연구 지원 사업 연구비와 2024년도 교육부의 재원으로 한국 연구 재단의 지원을 받아 수행된 지자체-대학 협력기반 지역 혁신 사업의 결과이다(2021RIS-003).

References

- 1) M. Haris and A. Glowacz, "RETRACTED: Road Object Detection: A Comparative Study of Deep Learning-Based Algorithms," *Electronics*, Vol.10, No.16, Paper No.1932, 2021.
- 2) W. Shin, D. Choi, H. Park and J. Kim, "Road Object Detection Robust to Distorted Objects at the Edge Regions of Images," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.7245-7251, 2024.
- 3) S. Kim, Y. Kim, H. Jeon, D. Kum and K. Lee, "Autonomous Driving Technology Trend and Future Outlook: Powered by Artificial Intelligence," *Transactions of KSAE*, Vol.30, No.10, pp.819-830, 2022.
- 4) M. Sun, D. Paek and S. Kong, "A Study on Deep Learning-Based Lidar Object Detection Neural Networks for Autonomous Driving," *Transactions of KSAE*, Vol.30, No.8, pp.635-647, 2022.
- 5) M. Iman, H. R. Arabnia and K. Rasheed, "A Review of Deep Transfer Learning and Recent Advancements," *Technologies*, Vol.11, No.2, pp.40, 2023.
- 6) J. Chen, D. Tam, C. Raffel, M. Bansal and D. Yang, "An Empirical Survey of Data Augmentation for Limited Data Learning in NLP," *Transactions of the Association for Computational Linguistics*, Vol.11, pp.191-211, 2023.
- 7) C. Shorten and T. M. Khoshgoftaar, "A Survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, Vol.6, No.1, pp.1-48, 2019.
- 8) R. Komatsu and K. Yamazaki, "Conditional Drive Environment Translation Using StarGAN with CBIN," *Proceedings of the 2021 4th Artificial Intelligence and Cloud Computing Conference*, pp.53-61, 2021.
- 9) B. Li, H. Guo and Z. Wang, "Data Augmentation Using CycleGAN-Based Methods for Automatic Bridge Crack Detection," *Structures*, Paper No. 106321, 2024.
- 10) W. Choi, J. Heo and C. Ahn, "Development of Road Surface Detection Algorithm Using CycleGAN-Augmented Dataset," *Sensors*, Vol.21, No.22, Paper No.7769, 2021.
- 11) Y. Peng, Z. Meng and L. Yang, "Image-to-Image Translation for Data Augmentation on Multimodal Medical Images," *IEICE Transactions on Information and Systems*, Vol.106, No.5, pp.686-696, 2023.
- 12) C. Ruffino, M. Prinet, M. Villanova and L. Martino, "Physically-Admissible Polarimetric Data Augmentation for Road-Scene Analysis," *Computer Vision and Image Understanding*, Vol.222, Paper No.103495, 2022.
- 13) K. Islam, M. Z. Zaheer, A. Mahmood and K. Nandakumar, "DiffuseMix: Label-Preserving Data Augmentation with Diffusion Models," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.27621-27630, 2024.
- 14) C. -M. Feng, K. Yu, Y. Liu, S. Khan and W. Zuo, "Diverse Data Augmentation with Diffusions for Effective Test-Time Prompt Tuning," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp.2704-2714, 2023.
- 15) B. Trabucco, K. Doherty, M. Gurinas and R. Salakhutdinov, "Effective Data Augmentation with Diffusion Models," *arXiv preprint, arXiv:2302.07944*, 2023.
- 16) B. Deng and Y. Lu, "Stable Diffusion for Data Augmentation in COCO and Weed Datasets," *arXiv preprint, arXiv:2312.03996*, 2023.
- 17) M. Morafah, M. Reisser, B. Lin and C. Louizos,

- “Stable Diffusion-Based Data Augmentation for Federated Learning with Non-IID Data,” arXiv preprint, arXiv:2405.07925, 2024.
- 18) X. Li, K. Kou and B. Zhao, “Weather GAN: Multi-Domain Weather Translation Using Generative Adversarial Networks,” arXiv preprint, arXiv:2103.05422, 2021.
 - 19) I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, “Generative Adversarial Nets,” Advances in Neural Information Processing Systems, Vol.27, 2014.
 - 20) J. Ho, A. Jain and P. Abbeel, “Denoising Diffusion Probabilistic Models,” Advances in Neural Information Processing Systems, Vol.33, pp.6840-6851, 2020.
 - 21) E. Sanchez and M. Valstar, “A Recurrent Cycle Consistency Loss for Progressive Face-to-Face Synthesis,” 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), pp.53-60, 2020.
 - 22) Y. Ma, X. Liu, L. Sun, J. Zhang, H. Wu and J. Wang, “Cycle Structure and Illumination Constrained GAN for Medical Image Enhancement,” Medical Image Computing and Computer-Assisted Intervention-MICCAI 2020: 23rd International Conference, Lima, Peru, Oct. 4-8, Proceedings, Part II, Vol.23, pp.667-677, 2020.
 - 23) T. Wang and Y. Lin, “CycleGAN with Better Cycles,” arXiv preprint, arXiv:2408.15374, 2024.
 - 24) Y. Zhao, R. Wu and H. Dong, “Unpaired Image-to-Image Translation Using Adversarial Consistency Loss,” Computer Vision-ECCV 2020: 16th European Conference, Glasgow, UK, Aug. 23-28, Proceedings, Part IX, Vol.16, pp.800-815, 2020.
 - 25) S. Lee and H. Leeghim, “Synthetic Infra-Red Image Dataset Generation by CycleGAN Based on SSIM Loss Function,” KIMST, Vol.25, No.5, pp.476-486, 2022.
 - 26) N. Kodali, J. Abernethy, J. Hays and Z. Kira, “On Convergence and Stability of GANs,” arXiv preprint, arXiv:1705.07215, 2017.
 - 27) Z. Wang, H. Zheng, P. He, W. Chen and M. Zhou, “Diffusion-GAN: Training GANs with Diffusion,” arXiv preprint, arXiv:2206.02262, 2022.
 - 28) T. Fang, R. Sun and A. Schwing, “DigGAN: Discriminator Gradient Gap Regularization for GAN Training with Limited Data,” Advances in Neural Information Processing Systems, Vol.35, pp.31782-31795, 2022.
 - 29) E. Becker, P. Pandit, S. Rangan and A. K. Fletcher, “Instability and Local Minima in GAN Training with Kernel Discriminators,” Advances in Neural Information Processing Systems, Vol.35, pp.20300-20312, 2022.
 - 30) X. Liu, Y. Zhang, H. Li and L. Wang, “LaDiffGAN: Training GANs with Diffusion Supervision in Latent Spaces,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.1115-1125, 2024.
 - 31) T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen and T. Aila, “Training Generative Adversarial Networks with Limited Data,” Advances in Neural Information Processing Systems, Vol.33, pp.12104-12114, 2020.
 - 32) R. Rombach, A. Blattmann, D. Lorenz, P. Esser and B. Ommer, “High-Resolution Image Synthesis with Latent Diffusion Models,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.10684-10695, 2022.
 - 33) C. Du, Y. Li, Z. Qiu and C. Xu, “Stable Diffusion Is Unstable,” Advances in Neural Information Processing Systems, Vol.36, 2024.
 - 34) L. Zhang, A. Rao and M. Agrawala, “Adding Conditional Control to Text-to-Image Diffusion Models,” Proceedings of the IEEE/CVF International Conference on Computer Vision, pp.3836-3847, 2023.
 - 35) J. -Y. Zhu, T. Park, P. Isola and A. A. Efros, “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks,” Proceedings of the IEEE International Conference on Computer Vision, pp.2223-2232, 2017.
 - 36) K. He, G. Gkioxari, P. Dollár and R. Girshick, “Mask R-CNN,” Proceedings of the IEEE International Conference on Computer Vision, pp.2961-2969, 2017.
 - 37) M. Arjovsky and L. Bottou, “Towards Principled Methods for Training Generative Adversarial Networks,” arXiv preprint, arXiv:1701.04862, 2017.
 - 38) S. Li, J. Wu, X. Xiao, F. Chao, X. Mao and R. Ji, “Revisiting Discriminator in GAN Compression: A Generator-Discriminator Cooperative Compression Scheme,” Advances in Neural Information Processing Systems, Vol.34, pp.28560-28572, 2021.