

시뮬레이션 기반 아모달 세그멘테이션 데이터 자동 생성 파이프라인 구축

이장훈, 하민세, 이종택*

경북대학교 IT대학 컴퓨터학부

tig06172@knu.ac.kr, haminse@knu.ac.kr, *jongtaeklee@knu.ac.kr

Simulation-Based Automatic Amodal Segmentation Data Generation Pipeline

Janghoon Lee, Minse Ha, Jong Taek Lee*

School of Computer Science and Engineering, Kyungpook National University.

요약

본 논문에서는 시뮬레이션 기반 아모달 세그멘테이션 데이터 자동 생성 파이프라인을 제안한다. 제안된 방식은 동일한 장면에 대해 가림막 유무 조건으로 두 차례 시뮬레이션을 수행하여, 가시 영역 마스크와 아모달 마스크를 각각 자동으로 생성하고, 이를 통해 가려진 영역의 마스크까지 포함한 정답 데이터를 구성한다. 해당 파이프라인은 수작업 라벨링 없이 대규모 고품질 학습 데이터를 효율적으로 구축할 수 있는 장점을 가진다. 이를 통해 다양한 가려짐 조건에서의 아모달 세그멘테이션 인식 성능 향상 및 후속 연구에 필요한 데이터 기반을 효과적으로 제공할 수 있을 것으로 기대된다.

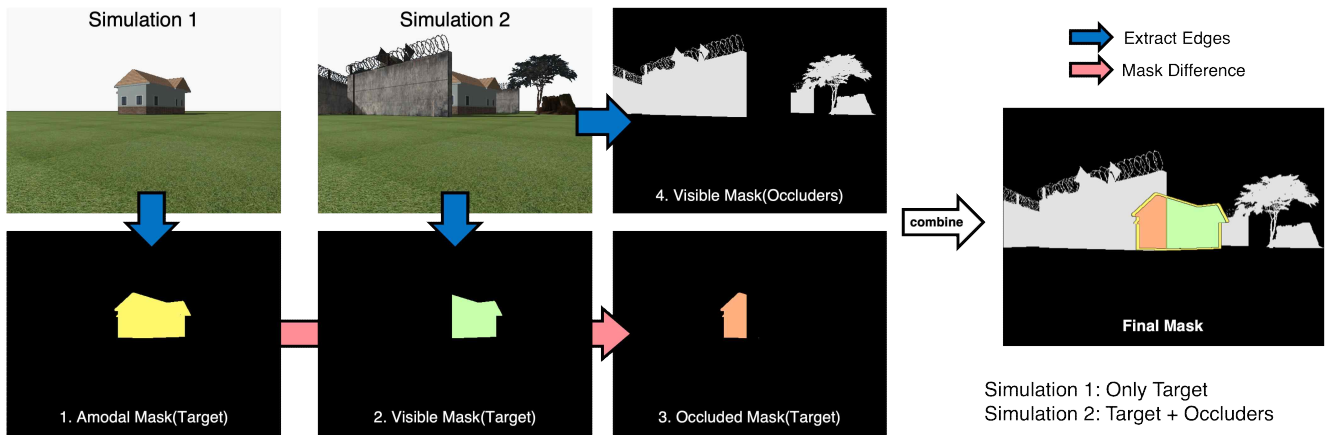


그림 1 아모달 세그멘테이션 데이터 생성 파이프라인

I. 서론

아모달 세그멘테이션은 객체가 다른 물체에 의해 가려져 보이지 않는 부분까지 포함하여 전체 객체를 세그멘테이션하는 작업이다 [1]. 실제 환경에서는 객체의 가려진 부분이 직접 관측되지 않기 때문에, 이러한 영역에 대해 라벨링을 수행할 때는 필연적으로 사람의 주관적 판단이 개입될 수밖에 없다. 따라서 수작업으로 아모달 세그멘테이션 데이터셋을 구축할 경우에는 많은 시간과 비용이 소요될 뿐 아니라, 사람에 따라 라벨링이 일관되지 않을 수 있어 데이터의 정확도와 일관성에도 한계가 존재한다. 이러한 주관적 오차와 비용 문제는 결과적으로 데이터셋의 품질 저하와 모델 성능 악화로 이어질 가능성이 높다. 더불어, 현재 존재하는 아모달 세그멘테이션 데이터셋은 그 규모가 제한적이고 특정 환경이나 객체에 편중되어 있어 다양한 상황과 객체를 포괄하는 데이터를 구축하는 데 어려움이 있다. 따라서 보다 객관적이고 대규모의 아모달 세그멘테이션 데이터셋을 효율적으로 구축할 수 있는 새로운 접근 방식이 필요하다. 본 연구에서는 이러한 문제를 해결하기 위해, 시뮬레이션 환경을 활용하여 객관적이고 정밀한 아모달 세그멘테이션 데이터셋을 생성하는 효율적인 방법을 제안하고자 한다.

II. 본론

본 연구에서는 Genesis 시뮬레이터를 기반 아모달 세그멘테이션 데이터셋 생성 파이프라인을 제안한다. 그림 1은 제안하는 데이터 생성 파이프라인에 대한 것이다 [2]. 이 파이프라인의 핵심은 고정된 카메라 위치에서 두 가지 장면을 렌더링하는 방식으로, 정확한 픽셀 수준의 라벨링을 자동으로 얻을 수 있도록 설계되었다. 하나는 가려짐 없이 대상 객체가 완전히 노출된 장면 (그림 1 - Simulation 1)이며, 다른 하나는 대상 객체가 다른 객체들에 의해 가려진 장면 (그림 1 - Simulation 2)이다. 두 장면에서 얻어진 라벨링 정보는 수작업 라벨링 과정에서 흔히 발생하는 주관적 오차와 비용 문제를 근본적으로 해결할 수 있다.

대상 객체가 완전히 노출된 장면에서는 대상 객체의 전체 형태를 명확히 나타내는 아모달 마스크를 얻으며, 이를 통해 대상 객체의 완전한 형태를 정확하게 확보할 수 있다. 반면 대상 객체가 가려짐이 발생한 장면에서는 실제로 관측 가능한 영역만을 나타내는 가시 영역 마스크 (visible mask)를 얻는다. 두 가지 마스크의 픽셀 단위 비교를 통해 명확히 가려진 영역을 정의하는 가려짐 영역 마스크 (occluded mask)가 생성되며, 이는 아모달 세그멘테이션의 핵심 정보가 된다. 또한, 가려짐이 발생한 장면에서 대상 객체의 마스크뿐만 아니라 대상 객체를 가리고 있는 객체들의 마

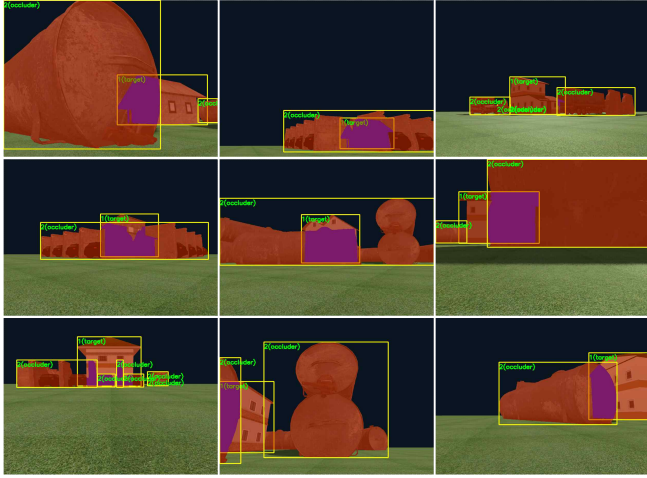


그림 2. 생성한 아모달 세그멘테이션 데이터 예제

스크린 가림막 마스크 (occluder mask)도 생성하여 각 객체의 상대적 위치와 가림 관계를 효과적으로 표현할 수 있다. 그림 2는 제안한 방법으로 생성한 데이터의 예제이며 붉은색과 보라색은 각각 가시 영역 마스크와 가려짐 영역 마스크를 나타낸다. 본 연구에서 생성된 데이터의 사례를 통해, 가려진 영역과 보이는 영역이 매우 명확하게 구분되며, 픽셀 단위의 정밀한 마스크와 명확한 인스턴스 레벨의 라벨링 및 바운딩 박스를 제공함으로써 데이터셋의 질적 우수성을 입증하였다. 특히, 제안된 파이프라인은 다양한 환경 및 객체 조건에 손쉽게 적용할 수 있는 확장성과 효율성을 가진다.

III. 결론

본 논문에서는 시뮬레이션 기반의 아모달 세그멘테이션 데이터셋 생성 방법을 소개하였다. 이 방법은 사람이 직접 생성하는 주관적이고 비용이 많이 드는 기존 방식에 비해 훨씬 정확하고 객관적이며, 효율적으로 대량의 데이터를 구축할 수 있는 장점을 가진다. 제안된 방법을 통해 생성된 데이터셋은 아모달 세그멘테이션 모델의 성능 향상과 기술 발전에 기여할 것으로 기대된다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 과학기술사업화진흥원의 지원을 받아 수행된 연구임(‘학연협력플랫폼구축 시범사업’ RS-2023-00304695)

참 고 문 헌

- [1] Li, Ke, and Jitendra Malik. "Amodal instance segmentation." Computer Vision - ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14. Springer International Publishing, 2016..
- [2] G. Authors, "Genesis: A universal and generative physics engine for robotics and beyond", (<https://github.com/Genesis-Embodied-AI/Genesis>), 2024