

가중치 기반 손실을 활용한 지능 온디바이스 통합 의미론적 분할 AI 모델 연구

김동연^{1*}, 조수현¹, 추인오¹, 윤만석¹, 권재익¹, 박태욱¹, 노승욱²

¹구미전자정보기술원, ²에이텍모빌리티

*dekim@geri.re.kr, shcho@geri.re.kr, iochu@geri.re.kr, yms@geri.re.kr, kwonjuk@geri.re.kr, ptu@geri.re.kr, sw_roh@atecmobility.com

A Study on an Intelligent On-Device Integrated Semantic Segmentation AI Model Utilizing Weighted Loss

Dong-Eon Kim^{1*}, Soo-Hyun Cho¹, In-Oh Chu¹, Mahn-Suk Yoon¹, Jae-Uk Kwon¹, Tae-Uk Park¹,
and Seong-Wook Roh²

¹Gumi Electronics & Information Technology Research Institute

²ATEC MOBILITY

요약

본 논문은 온디바이스 지능이 적용된 지능 로봇의 운용 환경에서 인지와 제어가 통합적으로 연동되는 구조를 제안한다. 효율적 작업 수행과 안전한 주행을 위해서는 강건한 장면 이해와 의미 분할(Semantic Segmentation) 기반의 환경 인식이 필수적이다. 그러나 YCOR, RELIS-3D, RUGD 등 공개 데이터셋은 각기 상이한 세부 클래스 정의를 포함하고 있어, 실제 온디바이스 운용 환경에 직접 적용하기 어렵다. 본 연구에서는 작업 가능 영역, 주행 가능 영역, 주행 불가 영역의 세 가지 통합 클래스 구조를 제안하여, 복잡한 분할 결과를 제어 가능한 형태로 단순화하였다. 또한 클래스 불균형 문제를 완화하기 위해 역빈도, 제곱근, 로그 기반의 가중치 손실 함수를 적용하고, DeepLabV3+ 모델을 이용하여 검증하였다. 제안된 접근법은 다양한 데이터셋을 일관된 구조로 재정의하여, 지능 로봇의 실시간 운용 신뢰성과 강화학습 기반 자율 제어 연구의 기반을 제공한다.

I. 서론

지능 로봇은 인간의 개입을 최소화하여 다양한 환경에서 자율적으로 작업을 수행할 수 있는 차세대 자동화 시스템으로, 비용 절감과 안전성 향상 측면에서 높은 잠재력을 지닌다. 이러한 시스템이 실제 환경에서 안정적으로 작동하기 위해서는 로봇이 주변 장면을 정밀하게 인식하고, 이를 바탕으로 주행·작업·회피 등의 의사결정을 수행할 수 있는 능력이 요구된다 [1].

의미 분할(Semantic Segmentation)은 이러한 인지-제어 연계를 위한 핵심 기술로, 픽셀 단위의 장면 이해를 통해 제어 신호로 직접 변환 가능한 정보를 제공한다 [2-4]. 그러나 기존의 공개 데이터셋은 각기 다른 연구 목적에 따라 세분화된 클래스 체계를 갖고 있어, 실제 현장에서의 적용성이 떨어진다. 따라서 본 연구에서는 온디바이스 기반 지능 로봇 운용을 위한 통합 의미 분할 체계를 제안하고, 단순하면서도 제어 연동성이 높은 3단계 구조를 정의하였다. 이를 통해 강화학습 기반 경로 계획, 작업 최적화, 안전성 확보 등 다양한 응용에서 활용 가능한 통합 인식 프레임워크를 구축하고자 한다.

II. 본론

2.1. 통합 의미 분할 클래스 체계(Unified Semantic Segmentation Classes)

본 연구에서는 기존의 세분화된 오프로드 데이터셋을 통합하여, 지능 로봇의 운용 요구에 직접 부합하는 세 가지 범주를 정의하였다.

표1. 로봇 작업용 클래스 분류 체계

구분	설명	주요 목적
작업 가능 영역	로봇의 작업 수행 대상이 되는 구역	작업 기능 활성화 및 목표 수행
주행 가능 영역	로봇의 이동이 가능한 영역	안전 이동 및 자율주행
주행 불가 영역	장애물, 위험 지역, 접근 금지 구역	회피 및 정지

이 통합 구조는 복잡한 클래스 구분을 제거하고, 인식 결과를 제어 로직에서 즉시 활용 가능한 형태로 단순화함으로써 실시간 의사결정 효율성을 높인다. 또한, 통합된 의미 지도는 강화학습 기반 경로 계획 및 작업 최적화 연구에서 일관된 상태 표현으로 활용될 수 있다.

2.2. 클래스 불균형 완화를 위한 가중치 손실 기법 (Weighted Loss for Class Imbalance)

의미 분할 학습에서는 데이터의 클래스 분포가 불균형할 경우, 희귀하지만 중요한 클래스의 인식 성능이 저하되는 문제가 발생한다. 이를 해결하기 위해 본 연구에서는 세 가지 가중치 기반 손실 함수를 적용하였다.

역빈도 기반(Inverse Frequency Weighting)은 희귀 클래스에 큰 가중치를 부여하여 적극적인 보정 효과를 제공한다.

$$w_c = \frac{total_{pxls}}{no_{classes} \times pxl_{count_c}} \quad (1)$$

여기서, w_c 는 클래스 c 의 가중치, $total\ pixels$ 는 전체 픽셀 수를 의미한다. 이 함수는 희귀 클래스에서 maximum emphasis를 제공하며 데이터 셋의 불균형이 심할 시에 효과가 있다.

제곱근 역빈도(Square Root Inverse Weighting)은 과도한 보정을 방지하면서 안정적 학습을 유도한다.

$$w_c = \sqrt{\frac{total_{pxls}}{class_{pxl_{count_c}}}} \quad (2)$$

여기서, $class\ pixel\ count_c$ 는 클래스 c 에 속하는 픽셀 수를 의미한다.

또한, 이 함수는 중간 수준의 rebalancing을 제공하며, 제곱근 함수를 사용함으로써 역빈도보다 완만한 조정을 제공한다.

로그 역빈도(Logarithmic Inverse Weighting)은 완만한 보정으로 클래스 간 분포를 자연스럽게 유지한다.

$$w_c = \log\left(\frac{total_{pxls}}{class_{pxl_{count_c}}}\right) \quad (3)$$

이 함수는, 로그 함수를 적용하여, 역빈도 대비 가장 완만한 보정을 제공한다.

이 세 가지 방법은 각 클래스의 픽셀 빈도에 기반하여 가중치를 산정하였으며, 학습 과정에서 희귀 클래스의 손실을 보강함으로써 전반적인 분류 균형을 개선하였다. 특히 제곱근 기반 접근은 학습 안정성과 일반화 성능 측면에서 가장 균형 잡힌 결과를 보였다.

2.3. DeepLabV3+ 기반 인식 모델 구성 (DeepLabV3+ Architecture and Configuration)

본 연구에서 활용된 DeepLabV3+ 모델은 인코더 - 디코더(Encoder - Decoder) 구조를 기반으로 하며, Atrous Convolution을 통해 다양한 수용 영역(Receptive Field)을 포괄하고, 정밀한 경계 복원을 위한 효율적 디코더 구조를 결합하였다. 모델의 백본(Backbone)은 ResNet-50 기반의 사전 학습(Pre-trained) 모델을 사용하여 특징 추출의 일반화 성능을 확보하였으며, 정규화(Normalization) 과정에서는 다중 GPU 학습 환경에서 일관된 통계량을 유지하기 위해 Synchronized Batch Normalization을 적용하였다. 또한 디코더 헤드(Decoder Head)는 다중 스케일 팽창(Atrous) 구조의 Depthwise Separable ASPP(Atrous Spatial Pyramid Pooling)을 구성하여 다양한 해상도에서의 문맥 정보를 효율적으로 통합하였고, 보조 헤드(Auxiliary Head)는 FCN(Fully Convolutional Network) 기반 보조 예측 모듈을 통해 학습의 안정성과 수렴 속도를 향상시켰다. 학습 정책은 SGD(Stochastic Gradient Descent) 옵티마이저와 PolyLR 학습률 스케줄을 사용하여 최적화하였으며, 이러한 구성은 야외 환경과 같이 복잡한 장면을 인식해야 하는 지능 로봇 운용 환경에서 경계 정밀도와 실시간성을 동시에 확보할 수 있도록 설계되었다.

2.3. 결과 및 분석

가중치 손실 기법을 적용한 학습 결과, 모든 방법이 기준선 대비 전반적인 인식 성능을 향상시켰으며 특히 제곱근 기반 접근이 가장 일관되고 안정적인 개선을 보였다. 이는 희귀 클래스에 대한 과도한 보정 없이 균형

잡힌 학습을 가능하게 하였기 때문이다.

시각적 검증 결과에서도 통합된 클래스 체계가 실제 장면에서 의미적 일관성을 유지하며, 로봇 제어 단계에서 즉시 활용 가능한 분할 결과를 생성하는 것이 확인되었다. 이는 복잡한 후처리 과정을 최소화하고, 강화학습 기반 제어 정책 개발에 직접 활용 가능한 인식 출력을 제공한다.

III. 결론

본 연구에서는 지능 로봇 운용 환경에서 인지 - 제어 통합을 실현하기 위한 3단계 통합 의미 분할 체계를 제안하였다. 제시된 구조는 기존 데이터 셋 간의 비일관성을 해소하고, 제어 로직과 직접 연결되는 단순화된 인식 출력을 제공함으로써 실시간 운용 효율성을 높였다. 또한 클래스 불균형을 완화하기 위한 가중치 손실 기법을 적용하여, 희귀 클래스에서도 안정적인 인식이 가능함을 확인하였다.

제안된 통합 프레임워크는 향후 강화학습 기반 자율 제어, 커버리지 최적화, 경로 계획 연구의 공통 인식 인터페이스로 활용될 수 있으며, 온디바이스 지능 로봇의 실시간 판단 및 안전 제어 시스템 구현에 기여할 수 있다.

ACKNOWLEDGMENT

이 논문은 2025년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구 결과 임(No. RS-2025-252455300, 지능 온디바이스 망연동 시험 플랫폼 개발).

참 고 문 헌

- [1] 노동희, 이해민, and 황근별. "과수원 환경에서의 농업용 로봇의 강건한 자율주행을 위한 디지털 트윈 기반의 관제 운영플랫폼 설계 및 구현." 한국통신학회 학술대회논문집 (2025): 879-880.
- [2] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
- [3] Dirgantara, Fussy Mentari, and Soo Young Shin. "Performance Analysis of Semantic Segmentation with YOLOv8: A Robust Approach for Autonomous Vehicle." 한국통신학회 학술대회논문집 (2024): 1081-1082.
- [4] 이도희, 이진혁, 김동우, & 신용구. (2022). Semantic segmentation 성능 향상을 위한 병렬 어텐션 모듈. 한국통신학회 학술대회논문집, 141-142.