

자율주행 대응 기계학습 데이터를 관리하고 분석하는 소프트웨어의 개발

*박종빈 **이한덕 ***김경원 ****정종진

전자부품연구원

*jpark@keti.re.kr

Development of Data Management and Analysis Software for Autonomous Vehicle Driving Environment

*Jongbin Park **Han-Duck Lee ***Kyung-Won Kim ****Jong-Jin Jung

Korea Electronics Technology Institute

요약

최근 기계학습 기술의 급속한 발전에 힘입어 자율주행을 위한 객체 인식 및 처리 기술 역시 비약적으로 발전하고 있다. 그러나 이러한 기계학습의 성능은 모델의 구조와 학습용 데이터의 품질에 영향을 받는다. 특히 주행환경을 잘 표현하는 학습 데이터가 중요한데 전혀 새로운 도로, 주행환경, 장애물, 정적 혹은 동적 객체 등을 마주하면 정확도와 안정성에서 부정적인 영향을 받을 수 있는 것이다. 해외의 주행 데이터들에 크게 의존하고 있는 우리나라의 현실에 비춰 볼 때 국내 환경에 맞는 학습데이터를 쉽고 효율적으로 확보/관리/분석할 수 있게 하는 환경의 구축이 시급하다. 따라서 본 논문에서는 자율주행을 위한 기계학습 데이터를 효과적으로 관리하고 분석하기 위한 소프트웨어를 설계하고 개발하였다. 구체적으로는 수집된 영상들을 관리하는 기능, 영상에 존재하는 노이즈 제거 및 화질 개선 처리 기능, 학습 및 검증을 위한 메타 정보 태깅 기능, 태깅 정보의 통계적 분석 기능들을 포함한다. 개발한 소프트웨어는 우리나라에서 자체 촬영한 자율주행 학습 영상들에 대해 딥러닝 모델들을 학습하고 검증하는데 활용할 예정이다.

1. 서론

자율주행은 인간의 개입이 매우 적거나 혹은 없는 상황에서 차량 스스로 환경을 판단하고 운행하는 것을 뜻한다. 이런 요구사항은 기계 스스로 높은 안정성을 확보하면서 실시간 고속 처리를 수행해야 하는 등 오랫동안 연구되고 있는 도전적인 분야이다. 그러나 최근의 각종 기술발전(임베디드 시스템의 고성능화, 광범위하고도 정확한 지도 정보의 확보, 각종 센서들의 정확도 향상, 딥러닝과 같은 기계학습 기술의 발전 등)으로 인해 제한된 공간에서의 자율주행은 이미 현실이 되었다 [1].

그러나 학습에 활용되지 않은 전혀 새로운 환경이나 장애물이 나타나면 예상하지 못한 오류를 일으킬 가능성이 생길 수 있으며 제한된 공간에서 확보된 학습데이터는 해당 학습 데이터 셋에 특화된 결과를 출력할 수 있다[2]. 이는 과적합(overfitting) 특성으로 불리기도 하는데 해결책으로 학습 중에 배치 정규화(batch normalization), 데이터 증강(data augmentation), 드랍 아웃(drop out) 등의 기술들을 적용할 수 있으나 보다 근본적인 해결책은 다양한 주행환경을 잘 표현하는 학습 및 검증 데이터를 확보하는 일이다 [3],[4].

구체적으로 주행 차량은 카메라 및 센서에서 수집된 위치정보, 주변 환경 정보, 차량 움직임을 포함하는 센싱 정보 등을 확보하게 되는데 이런 데이터들을 사용하여 주행에 필요한 세부적인 규칙들을 자동으로 학습하는 것이 자율주행을 위한 딥러닝 기술의 핵심이다. 하지만 인식을 위한 학습 및 검증 데이터는 주로 해외에서 수집되고 배포된 자료에 의존하는 실정이고 이는 국내의 도로 환경을 잘 반영하지 못하는

문제가 있다. 일례로 우리나라에만 존재하는 도로정보 표지판, 속도제한 표지판, 신호등, 차량 번호판, 차종, 도로 주변 객체 등은 해외 데이터에서 당연히 제공하지 않는다. 이런 문제를 해결하고자 본 논문에서는 수집된 데이터의 관리, 화질 처리, 메타정보 태깅, 태깅 정보의 분석을 통합적으로 수행하는 소프트웨어를 개발하였다.

2. 본론



그림 1. 수집 가능한 주행환경 정보와 개발한 소프트웨어의 요소모듈

그림 1은 차량에서 수집할 수 있는 주행환경 정보와 본 논문에서 개발한 소프트웨어 요소 모듈들을 나타낸다. 그림 1과 같이 주행차량은 1개 이상의 다시점 카메라로부터 영상정보, 라이다(LIDAR, Light Detection and Ranging)센서로부터 3차원 거리정보, GPS(Global Positioning System)정보와 같은 지도상에서의 위치정보, 차량 내부의 움직임이나 각종 상태를 샘플링한 센싱정보, 차량 인근의 장애물 정보

등을 수집할 수 있다. 이런 정보들은 통상적으로 시간 태그가 함께 저장되므로 수집 당시 장치 간 시간 동기화가 잘 되어 있다면 독립적으로 수집했다더라도 나중에 동기화 분석이 가능하다. 개발한 소프트웨어는 이렇게 차량에서 수집된 정보들을 데이터베이스에 저장 및 관리하고, 해당 데이터에 메타정보를 태깅하고, 태깅된 정보를 기반으로 통계적 분석을 수행하며, 데이터 상호간 연관성 분석 및 검색 및 추천을 위한 신호처리를 수행하도록 구현했다.

2.1 구현내용

그림 2, 3, 4, 5는 상기 소개한 기술들을 구현한 소프트웨어의 사용자 인터페이스로써 데이터베이스(DB) 기반으로 수집 데이터를 관리하고, 화질 개선 처리 및 각종 영상 편집, 영역을 설정하고 태깅하기, 태깅 정보를 기초로 통계 분석을 수행하는 모습을 보여준다. 사용자 인터페이스는 C# 언어를 사용하여 구현했으며, 영상 데이터의 관리를 위해서는 Maria DB, 정보 태깅을 위해서는 Mongo DB를 사용했다. 학습을 위해 수집된 데이터는 통상 폴더 단위로 위치하기 때문에 별도의 프로젝트 파일 없이도 영상 데이터 셋을 폴더 단위로 그룹화하면 이를 자동으로 DB에 등록하도록 했다. 태깅 정보 처리는 [5]에서 배포한 공개 소스코드를 일부 수정하여 소프트웨어 집적했다. 본 구현의 특징으로는 태깅된 정보들을 기반으로 태그들의 발생 빈도를 시각화하고, 태그들 사이의 계층적인 스카마 구조를 고려하여 학습 데이터 셋의 특성 분석이 가능하도록 했다는 점이다.

3. 결론

본 논문에서는 자율주행을 위한 학습 데이터의 관리/분석을 위한 통합 소프트웨어를 개발하였다. 구현한 기능 중에서 수집된 영상들을 관리하는 기능, 영상에 존재하는 노이즈 제거 및 화질 개선 처리 기능, 학습 및 검증을 위한 메타 정보 태깅 기능, 태깅 정보의 통계적 분석 기능들이 서로 유기적으로 연결되도록하여 실제 메타정보 추가 및 분석에 유용하게 쓰이도록 하였다.

감사의 글

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No. 2019-0-00399, 비정형 주행 환경 대응이 가능한 자율차 탑재용 AI 기반 인지, 판단 및 제어 솔루션 개발)

참고문헌

- [1] M. Bojarski and et al., "End to end learning for self-driving cars", arXiv preprint arXiv:1604.07316, 2016.
- [2] R. Sebastian Ramos and et al., "Detecting unexpected obstacles for self-driving cars: Fusing deep learning and geometric modeling", 2017 IEEE Intelligent Vehicles Symposium , 2017.
- [3] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning", arXiv preprint arXiv:1712.04621, 2017
- [4] Y. LeCun, and Y. Bengio and J. Hinton, "Deep learning", Nature, vol. 251, no 7553, pp.436-444, 2015.
- [5] Semantic Segmentation Editor, URL: <https://github.com/Hitachi-Automotive-And-Industry-Lab/semantic-segmentation-editor>

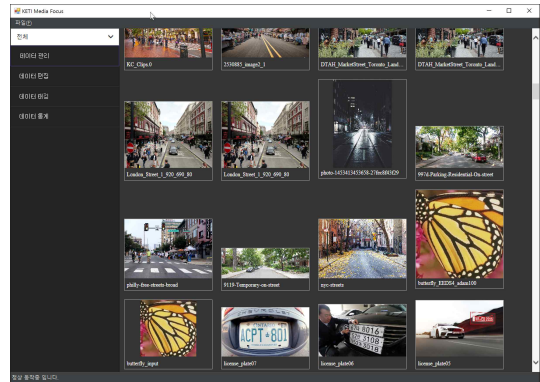


그림 2. 학습 데이터 태깅을 위한 미디어 관리

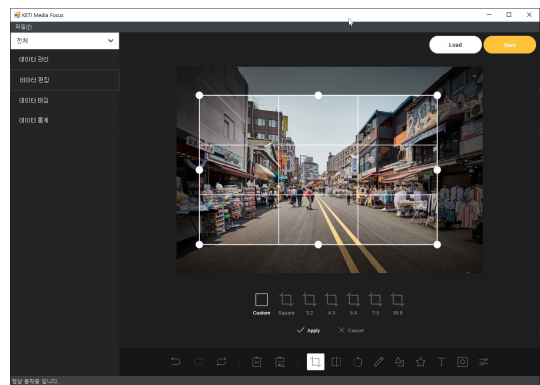


그림 3. 영상 처리 및 크기 조절

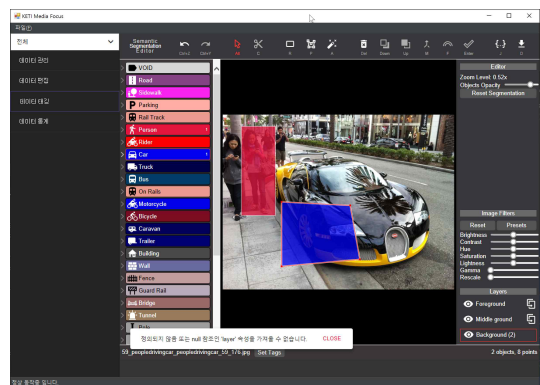


그림 4. 학습 데이터 태깅 UI ([5]기반 구현)

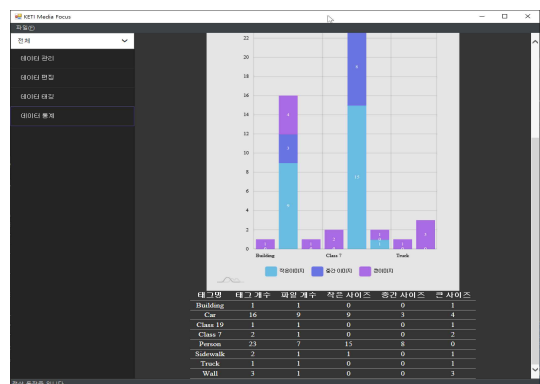


그림 5. 학습용 데이터의 통계 분석 UI