

# 실내 문서 전달 자율주행 카트 개발을 위한

## 객체 인지 및 충돌 방지 모듈 개발

### An Implementation of Object Recognition and Collision Avoidance Modules for Indoor Self-driving Cart

#### 요 약

최근 머신러닝 기술이 자동차에 접목 되면서 ‘자율주행’에 관한 연구가 활발히 이루어지고 있다. 관련 기술은 실내에서도 충분히 적용가능하기 때문에 본 논문은 실내문서전달용 자율주행 카트 개발을 위해 기계학습 기반의 실시간 객체인지 및 충돌 방지 모듈 개발을 목표로 한다. 최종적으로 수집된 실내정보를 클라우드에서 분석 후 주행을 위한 정확한 실내정보를 인지하여 안정적으로 실내를 주행하는 자율주행 카트 개발을 목표한다.

#### 1. 서 론

##### 1.1 연구 배경

5G network 환경과 더불어 머신 러닝과 같은 SW기술발전으로 인해 점차적으로 자율 주행 기술의 점진적 변화가 예상된다. 미국 자동차 학회(SAE)는 자율주행의 구현 정도에 따라 자율주행 기준 발전단계를 마련하였고, 현대자동차는 2022년 까지 주행관련 정보를 판단하고 행동을 결정할 수 있는 레벨3단계를 대부분 구축할 것으로 전망하고 있다. 이처럼, 실외용 Autonomous Vehicle의 연구가 현재 활발히 연구 개발되고 있으나, 도로 위가 아닌 실내의 경우 또 다른 문제점에 직면하게 된다. 먼저 Guide Line인 차선이 없다는 점과 더불어, 비좁은 통로가 존재하고 사람들의 왕래가 실내에서는 더욱 잦다는 점을 감안하면 단순히 위치정보만을 활용한 자율주행카트 구현에는 무리가 있다. 따라서 본 연구는 기존 실외 자율주행카트의 개념에 덧붙여 실시간 영상처리, 머신러닝을 통한 Object Recognition and Collision Avoidance Modules을 구현하고 정확한 실내정보를 인지하여 안정적인 실내 자율주행 카트 개발을 목표한다.

##### 1.2 연구 개발 목표

실내 자율주행 카트의 위치정보는 클라우드에 저장되고, 이와 동시에 전면, 측면에 대한 실시간 영상처리를 통해 보다 정확한 위치정보를 얻을 수 있도록 자율 주행 카트 모듈을 구현한다. 이에 덧붙여, 실내에서의 다양한 텍스트(실내의 각 태그 혹은 문패)의 특징점을 기반으로 위치정보를 특정한다. 또한, 영상처리 모듈을 통해 실시간 객체(사람) 인식 기능을 구현하여 1)Human Oriented Driving System을 구현한다.

#### 2. 관련 연구

##### 2.1 SLAM(Simultaneous Localization And Mapping) 기반 자율주행 카트

초음파, 적외선, 관성 측정 장치인 IMU와 LIDAR, 모터 등을 사용하여 제어 가능한 로봇에 대해연구가 이루어진 바 있다. 이 원격제어 로봇은 LIDAR센서를 통해 물체와의

거리를 측정하여 장애물 회피기능을 구현하였다. 다만 2)LIDAR센서가 Object의 반사펄스를 이용하여 물체의 거리를 측정하는 특성을 고려하면, 여전히 반사율이 높은 환경에서 정확한 거리인식이 어렵고, 이에 따른 우발적 장애요소에 대한 회피문제점이 잔존한다.

##### 2.2 클라우드 연계 자율주행 맵 시스템 기술 동향

클라우드와 연계하여 생성되는 자율주행 맵을 구성하면 보다 정밀성 있는 맵을 구성할 수 있다. 또한 도로 교통과 관련된 실시간 갱신주기 측면에서 더욱 세밀하고 정확한 위치정보를 제공할 수 있다. 그러나 실내에서의 경우 주행 경로 상의 장애물이 있을 가능성이 높고, 좁은 통로의 특성상 많은 우발적 요소가 예상되기 때문에 정적인 지도만을 구축하여 주행하기에는 한계점이 있다.

##### 2.3 위치 추정, 충돌 회피, 동작 계획이 융합된 이동 로봇의 자율 주행 기술

실내 이동로봇의 자율주행 방법으로 지도생성, 위치 추정, 장애물 회피, 경로 계획에 대한연구가 이루어진 바 있다. 위치 추정을 위해서 지도 정보를 이용, 센서 데이터를 계산하였고 인공지능위계를 사용하여 장애물로부터의 척력, 목표위치로의 인력을 구하여 3)회피 알고리즘을 구현하였다. 그러나 4)LRF센서를 바탕으로 장애물을 인식하였고 때문에 실내의 좁은 통로 특성상센서 길이의 제약으로 인해 장애물 감지거리가 55cm에 불과하였다.

##### 2.4 Multi-digit Number Recognition

CNN을 이용해 이미지에서 여러 자리의 숫자를 감지하는 5)연구가 존재하며, 해당 연구에서 사용된 SVHN(Street View House Numbers) 6)데이터셋을 훈련 용도로 사용하고자 한다.

2) 레이저 펄스의 반사 시점까지의 시간을 측정하여 반사체의 위치좌표를 측정하는 센서.

3) 해당 연구에서는 이동로봇의 위치추정방법으로 MCL알고리즘, 삼각측량, 칼만 필터 등을 사용하였음.

4) LIDAR 센서와 동일한 원리로, 레이저펄스를 이용하여 반사체간의 거리를 측정하는 센서.

5) 참고문헌[5] 참조

6) <http://ufldl.stanford.edu/housenumbers>

1) 이동객체(사람) 회피 주행

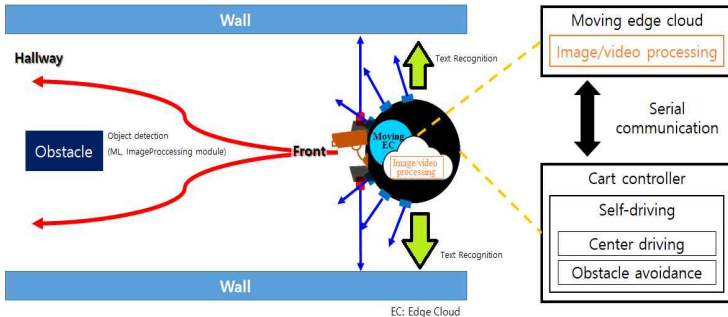
## 2.5 기존 연구 분석

[표 1] 기존연구 분석

구분	연구내용	한계점
SLAM기반 자율주행 카트	IMU, LIDAR, 모터를 사용하여 자율주행 구현	반사펄스를 이용하는 LIDAR, 초음파센서의 특성상, 반사율이 높은 환경에서 우발적 상황에 대한 한계점 잔존
클라우드 연계 자율주행 맵 시스템 기술동향	클라우드 기반 지도 구축, 엣지클라우드 - 카트 간 상호작용을 통한 자율주행 구현	실내의 경우 주행 경로 상에 장애물이 있을 가능성과 좁은 도로 폭을 고려해 볼 때 정적인 지도만을 바탕으로 한 자율주행에는 많은 어려움이 존재함.
위치 추정, 충돌 회피, 동작 계획이 융합된 이동 로봇의 자율주행 기술	LRF 센서를 바탕으로 한 장애물 인식, 인공지능위계를 사용하여 회피알고리즘 구현	실내의 좁은 통로 특성상 LRF센서의 길이가 짧아짐에 따라 장애물 감지거리가 약 55cm에 불과하여 실질적인 자율주행에는 어려움이 있음.

앞서 보듯, 관련 기존 연구는 정적인 맵을 구축하거나, 센서에만 의존한 주행으로 안정성의 한계가 있었다. 따라서 본 연구는 LIDAR센서를 통한 근거리 감지와 동시에 실시간 영상처리를 통한 원거리 이동 객체(사람) 인지, 실내 특징점(Door Plate Number)인식을 바탕으로 실내 환경에 특화된 보다 효율적인 자율주행카트를 제시한다.

## 3. 주행 시나리오 및 모듈별 기능



[그림 1] 주행 개념도(위)

[표 2] 주행 개념도 기호

구분	상세 내용	구분	상세 내용
→	추후 카트 예상 이동 경로(장애물 회피)	↓	카트 양 측면 Door Plate의 Number인식
↑	LIDAR센서를 통한 근거리 객체 감지	Front	카트 전면 부 이동 객체(사람) 인지

### 1) 기기구성

자율주행 카트에 LIDAR센서를 전방, 양 측면에 부착하여 근거리에 대한 물체감지가 이루어진다. 또한, 카트의 위치 정보는 각 클라우드의 엣지에 의해 세밀화 된다. 다만 클라우드의 엣지만으로는 카트의 실내위치를 정확히 추론할 수 없음과 LIDAR센서의 감지거리한계를 고려하여, 다음 두 모듈을 통해 실내의 위치를 정확하게 추론하고 원거리의 물체를 회피할 수 있도록 한다.

### 2) 영상처리모듈

영상처리 모듈은 카트에 부착된 전면카메라와 양 측면 카메라를 활용한다. Object Recognition 모듈은 전면 카메라의

실시간 영상으로부터 객체를 회피한 예상주행경로를 제어모듈로 전송한다. 다음으로 양 측면 카메라는 실내의 정보파악을 위한 Door Plate Number 추출 기능을 위해 동작한다. Number Tag Detection 모듈에 의해 영상의 각 프레임마다 실내의 Door Plate에 대해 Number Tag를 크롭하며 이후 크롭된 이미지의 해석을 위해 동일 라즈베리파이애 로드된 머신러닝 모델로 해당 이미지를 반환한다.

### 3) 머신러닝 모듈

실내 환경에 특화된 모델에 의해 크롭 이미지의 노이즈를 필터링한다. 또한 정확한 Number Tag값을 해석하여 보다 정확한 실내의 위치를 추론한다. 최종적으로, 주행에 관한 정보가 도출되고 해당 정보를 카트의 제어모듈로 전송, 이를 기반으로 안정적인 실내주행을 구현한다.

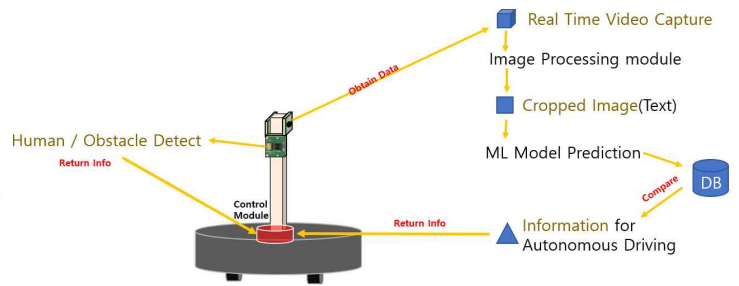
### 3.1 구현 환경

Object Recognition and Collision Avoidance 모듈은 라즈베리파이4B모델에 별도 카메라 모듈을 부착하여 구성되었으며 라즈베리파이의 하드웨어 사양은 다음과 같다. 또한, 카트의 동작 범위는 좁은 복도가 있고, 명패(Door Plate Number)가 있는 실내 환경으로 한정한다.

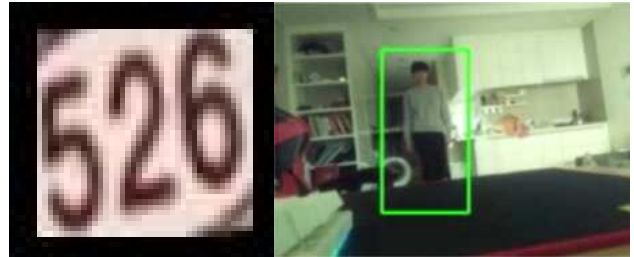
[표 3] 주요 모듈 개발 환경

구분	구현 환경
	[라즈베리파이 4B모델] 브로드컴 BCM2711, 쿼드코어 Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz 4GB LPDDR4-3200 SDRAM 카메라모듈 V2 8Megapixel

### 3.2 세부 시나리오



[그림 2] 모듈간 데이터 처리과정



[그림 3, 4] 실내 태그 크롭이미지(좌) 및 사람 인식(우)

· Door plate Number Detection 시나리오 : 양 측면 카메라에서 전달받은 실시간 RGB 영상에서 Door plate에 해당하는 이미지를 48\*48 사이즈로 크롭한다. 크롭된 이미지는 Grayscale로 변환되어 CNN 모델에 입력된 후, 모델은 6개의 결과 값을 도출한다. 이 때 6개의 결과 값은 이미지가 숫자를 포함하는지 여부, 인식된 숫자의 길이, 최대 4자리의 숫자 값으로, 이 결과 값을 통해 실내 정보 데이터베이스와 비교하여 현재 카트의 정확한 실내 위치를 카트의 제어모듈로 전송한다.

· Human Object Recognition시나리오 : 전방 카메라에서 받은 실시간 영상으로부터 OpenCV의 Hog디스크립터를 이용하여 사람을 인식한다. 이후 카트의 예상 동선을 10개요소를 가진 1차원배열로 작성하고 카트의 너비를 고려, 가장 이상적인 경로 정보를 1차원 배열 형태로 제어모듈에게 전송한다.

#### 4. 한계점 및 해결 방안

##### 4.1 인식 시간

현재 pytesseract의 OCR라이브러리는 720\*960 이미지 파일 기준, Text Extraction의 지연시간은 평균 0.38초로 측정되었다. 라즈베리파이의 경우 CPU, GPU환경이 좋지 않아 많은 Latency가 예상되므로 기존 OCR라이브러리 보다는 학습된 머신러닝 모델을 통하여 예측하는 것이 바람직할 것으로 보인다. 7)현재 라즈베리파이 환경에서의 실시간 Human Detection 실행시간은 약 0.05 - 0.06초로 실시간 Object인식에는 무리가 없다. 다만 Number Tag추출의 경우 더 높은 해상도와 머신러닝 모델의 지원이 필요한 부분이므로 학습모델의 경량화와 프레임 조절 및 정확도 높은 크롭(Crop)이 요구된다.

##### 4.2 객체 인지

Object Recognition의 경우 가장먼저 Human Oriented로 구현되어야 하며, 사람 중심으로 철저히 구동되어야 한다. 최소 5m에서 최대 25m거리에서 사람을 인식할 수 있어야 하며, 우발적인 동작에도 빠르게 반응할 수 있어야 한다.

##### 4.3 실내 Door Plate Number Tag 정보 인식

Number Detection의 경우 라즈베리파이의 카메라모듈 프레임이 초당 20-30프레임을 감안하여 최소 4-5%이상의 이미지 크롭 정확도를 요구한다. 하지만 이는 최소범위이므로, 원활한 주행을 위해서는 약 40%이상의 이미지 크롭 정확도를 요구한다. 머신러닝 모델의 경우 약 80%이상의 정확도를 유지하여, 실내 환경에서의 Text를 예측하고 8)잡음을 걸러낼 수 있어야 한다.

##### 4.4 Door Plate Number 인식용 데이터셋 활용

Number Detection의 용도로 사용될 데이터셋은 3~4 자리의 숫자로 이루어진 이미지 집합을 요구한다. 이 경우 기존 연구에서 사용되었던 SVHN 데이터셋만으로는 높은 정확도를 기대하기 어려우므로 기존 SVHN 데이터셋에 실내 환경에 특화된 데이터를 포함하여 데이터셋을 새로 구축할 필요가 있다.

#### 5. 성능평가

실내조명의 영향과 Door Plate의 외형이 제각각 상이함에 따라 개발 초기, 이미지 크롭의 잡음으로 인한 정확한 실내정보 파악이 불가능 하였다. 때문에 여러 크롭 전략을 마련하여 100장의 이미지를 대상으로 9)Door Plate Detection실험을 진행하였으며 최종적으로 크롭이미지 정확도를 84%까지 개선하였다.

##### 전략1) Gaussian Blur 적용

DoorPlate의 작은 윤곽선이 Blur처리 되어 정확한 크롭이 불가능하였다. 크롭 정확도는 8%에 그쳤으며, 나머지 92%의 이미지가 잡음이었다.

7) 320\*240 해상도, Raspberry pi 4모델 기준으로 측정.

8) 실제 Text정보가 아닌, 반사율과 같은 여러 외적요인에 의해 잘못 크롭된 이미지

9) 동일 영상에서 Door Plate를 대상으로 크롭의 성능 테스트를 진행하였음.

정확도는 크롭이미지 100장 중 올바른 Crop이 이루어진 이미지의 비율로 규정함

##### 전략2) Gaussian Blur미적용, 크롭이미지 y좌표 150픽셀 이상의 잡음 배제

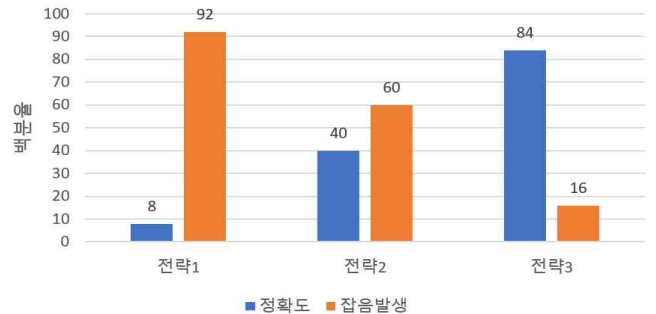
Door Plate가 영상의 상단에 위치하고 있는 점을 고려하여 특정 영역에서만 Detection이 이루어지도록 구성하였다. 정확도는 약 42%에 달했으나 여전히 전체 크롭 이미지 중 대부분이 잡음이었다.

##### 전략3) 크롭된 이미지의 평균 RGB값 추출, 범위 만족 여부 확인

기존 크롭 전략 2번에 덧붙여 크롭된 이미지를 10\*10 픽셀크기의 이미지로 resize하였다. 이후 해당 이미지의 전체 평균 RGB색상을 추출하여, 기존 크롭 이미지들의 평균 RGB색상과 비교하여, 색 범위 만족여부를 확인하였다. 정확도는 약 84%로, 보다 정확한 Door Plate Detection이 가능하였다.

[표 4] Door Plate Number Detection 실험

구분	주요 특징	정확도	잡음 발생
전략1	Gaussian Blur(3,3)적용	8%	92%
전략2	Gaussian Blur(3,3)미적용 크롭이미지의 y좌표 150 픽셀 이상 잡음 배제	40%	60%
전략3	크롭된 이미지의 평균 RGB값 추출, 필터링 적용	84%	16%



[그림 5] 크롭 전략별 정확도 및 잡음발생도 측정 결과

#### 6. 결론 및 향후 연구

실내용 자율 주행 카트를 개발함에 있어서, 이미 저장된 지도정보가 아닌, 상황에 따라 유연한 대응을 하는 자율주행카트를 개발하고, 실내 환경에 최적화된 자율주행 시스템을 연구하였다. 또한 근거리의 경우 센서를 활용한 Object Detection과, 원거리의 경우 영상처리 및 머신러닝 모델에 의한 예측으로 우발적인 상황에는 하드웨어적인 명령으로 빠른 대처가 가능하며 실시간 영상처리를 기반으로 향후 이동경로를 미리 예상할 수 있음을 감안하면 정적 로드맵을 구축하는 것보다 훨씬 효율적인 주행을 기대할 수 있다. 향후 연구로는 실제 문서전달 기능을 감안하여 하드웨어 및 소프트웨어적으로 최적화된 기기를 구현할 계획이다.

#### 참 고 문 헌

- [1] 김용년, and 서일홍. “바닥 특징점을 사용하는 실내용 정밀 고속 자율 주행 로봇을 위한 싱글보드 컴퓨터 솔루션.” 로봇학회논문지 14.4 (2019): 293-300.
- [2] 박세준, 서용호, and 양태규. “위치센서와 레이저거리센서를 이용한 실내용 이동로봇의SLAM.” 한국정보기술학회논문지 8.11 (2010): 61-69.
- [3] 노성우, 고낙용, 김태균, Noh, Sung-Woo, Ko, Nak-Yong, and Kim, Tae-Gyun. “위치 추정, 충돌 회피, 동작 계획이 융합된 이동 로봇의 자율주행 기술 구현.” 한국전자통신학회 논문지 = The Journal of the Korea Institute of Electronic Communication Sciences 6.1 (2011): 148-56. Web.
- [4] 최정단, 민경욱, 성경복, 한승준, 이동진, 박상현, 강정규, 조용우, Choi, J.D., Min, K.W., Sung.K.B., Han, S.J., Lee, D.J., Park, S.H., Kang, J.G., and Jo, Y.W. “클라우드 연계 자율주행 맵 시스템 기술동향.” 전자통신동향분석 = Electronics and Telecommunications Trends 32.4(2017): 40-47. Web.
- [5] Ian J. Goodfellow, Yaroslav Bulatov, Julian Ibraz, Sacha Arnoud, Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks, , Vinay Shet, 2011