

논문 2024-13-29

국내 도심에서 자율주행을 위한 신호등 인식 모듈 및 데이터 셋 구축 프로세스 설계

(Design of Building Dataset and Traffic Light Recognition Module for Domestic Urban Autonomous Driving)

박재형[†], 이진희[†], 김제석, 권순*
(Jaehyeong Park, Jin-Hee Lee, Je-Seok Kim, Soon Kwon)

Abstract : In the context of urban autonomous driving, where various types of traffic lights are encountered, traffic light recognition technology is of paramount importance. We have designed a high-performance traffic light recognition module tailored to scenarios encountered in domestic urban driving and devised a dataset construction process. In this paper, we focus on minimizing the camera's dependency to enhance traffic light recognition performance. The camera is used solely to distinguish the color information of traffic lights, while accurate location information of the traffic lights is obtained through localization and a map. Based on the information from these components, camera RoIs (Region of Interest) are extracted and transmitted to the embedded board. The transmitted images are then sent back to the main system for autonomous driving control. The processing time for one traffic light RoI averages 43.2 ms. We achieve processing times of average 93.4 ms through batch inference to meet real-time requirements. Additionally, we design a data construction process for collecting, refining, and storing traffic light datasets, including semi-annotation-based corrections.

Keywords : Traffic light, Embedded system, Autonomous driving

1. 서론

자율주행 차량을 개발하기 위해서는 객체 인지, 위치 추정, 경로 계획, 제어 등의 복합적인 기술을 요구한다. 특히, 다양한 신호등이 있는 국내 환경에서 도심 자율주행을 위해 신호등 인식 기술은 주변 차량뿐 아니라 보행자의 안전과도 직결되므로 고성능의 인식 기술이 필수적으로 요구된다. 이런 이유로 기존 연구자들은 신호등 인식 기술 개발의 중요성을 인지하면서도 완성도 높은 신호등 인식 기술 연구는 느리게 진행되고 있다. 따라서 이 논문에서는 도심 자율주행에서 반드시 필요한 신호등 인식 문제를 해결하기 위해, 자율주행에 필수 불가결 요소들 (맵, 센서)을 이용하여 고성능의 신호등 인식 모듈을 개발하였다. 대부분의 도심 자율주행 플랫폼은 신호등 인식을 위해 V2X와 같은 통신 환경을 구축하거나, 색상 정보를 활용할 수 있는 카메라 센서를 사용한다 [1, 2]. 하지만, 통신 기반 기술은 모든 신호등에 추가적인 인프라 구축이 필요하며, 카메라 기반 기술은 전체 이

미지에서 신호등을 검출하기 때문에 비교적 높은 연산이 요구되는 반면 인식 성능이 낮고, 자차 신호 여부 판단과 같은 추가 연산이 필요하다. 본 논문에서는 신호등 인식에서 카메라의 의존성을 최소화하기 위해, 로컬라이제이션 기술과 맵을 활용하여 신호등의 위치 정보로 카메라 이미지에서 신호등 RoI (Region of Interest)를 추출하여 인식 성능을 높이는 데 주력하였다. 우리의 자율주행 플랫폼은 정확한 객체 검출과 로컬라이제이션을 위해 라이다 센서를 이용하고, 경로 계획을 위한 맵을 활용하므로, 추가적인 비용 없이 제안하는 신호등 인식 모듈에 필요한 정보를 이용할 수 있고 주행 경로의 신호등만을 인식하므로 처리 시간 및 비용도 현저히 줄어들도록 설계하였다. 자율주행 플랫폼에 활용하기 위해 저전력 임베디드 시스템에서 구동하였다. 또한, 효율적인 신호등 데이터 셋 구축을 위한 자동화 프로세스를 제안한다. 기존의 관련 연구들과 우리가 제안하는 방법을 비교 설명하고, 3장에서는 타겟 시나리오 선정 및 센서 환경구축과 제안하는 신호등 인식 모듈의 아키텍처에 대해 설명하고, 4장에서는 효율적인 신호등 데이터 셋 구축을 위한 자동화 프로세스에 대해 설명한다. 5장에서는 타겟 시나리오에서 자차 구축한 신호등 데이터 셋과 공개 데이터 셋인 LISA로 모델간 정량평가 및 성능을 비교한다. 6장에서는 결론 및 향후 연구에 대해 설명한다.

[†] Equal Contribution (stillrunning, jhlee07@dgist.ac.kr)

* Corresponding Author (soonyk@dgist.ac.kr)

Received: Mar. 4, 2024, Revised: Apr. 14, 2024, Accepted: Jul. 26, 2024.

J. H. Park: DGIST(Researcher)

J. H. Lee, J. S. Kim: DGIST(Senior Researcher)

S. Kwon: DGIST(Principal Researcher)

※ 본 연구는 과학기술정보통신부에서 지원하는 대구경북과학기술원 기 관고유사업 (24-IT-01), 기술사업화 역량강화사업 (2023-DG-RD-0041-0 2-201), 연구개발특구진흥재단사업 (2023-TB-RD-0017-01) 지원을 받아 수행되었습니다.

II. 관련 연구

신호등 검출은 일반적인 객체 검출과 구별되는 몇 가지 특성이 있다. 이는 신호등의 다양한 크기, 배경의 높은 가변성, 종류 및 신호의 다양성이 국가나 지역마다 다르기 때문이다. 기존 연구에서는 이러한 문제를 해결하기 위해 다양한 접근 방법을 제안했으며, 이를 다음과 같이 크게 두 가지 그룹으로 분류할 수 있다. 기존 객체 검출기의 변형한 방식과 신호등 위치 파악 및 상태 분류 단계로 수행하는 다단계 접근 방식이다.

1. 객체 검출기의 변형

객체 검출기를 변형하는 접근 방법은 기존의 객체 검출기를 변형하여 작은 크기의 신호등을 인식하도록 한다. 이 접근법은 빠른 추론 속도를 확보하기 위해 대표적인 1단계 객체 검출기인 YOLO 계열 [3]의 변형된 모델을 주로 활용한다 [4, 5]. 예를 들어, Wang 등 [4]은 YOLOv4를 기반으로 가벼운 특징을 향상하는 기법과 Bounding box 불확실성 예측하는 방법을 추가하여 인식 성능을 향상시켰다.

이 방법들은 독립적으로 구동할 수 있으나, 자율주행에 활용하기 위해서는 자차 신호등 결정, 미인식된 신호등에 대한 대처와 딜레마 존에서의 대응 등의 추가적인 기능 구현이 요구된다.

2. 다단계 접근법

다단계 접근 방법은 신호등을 여러 하위 작업을 거쳐 인식하는 방식이다. Handcraft feature 또는 딥러닝 기반의 객체 검출기와 분류기를 조합하거나, 사전정보 기반의 검출기와 분류기를 조합하는 방식을 포함한다. Hirabayashi 등 [6]은 자율주행 프레임워크인 Autoware [7]를 활용하여 3D 맵에서 신호등 정보를 추출한 뒤, 이를 이미지로 투영하여 신호등 위치를 검출하였다. 검출된 영역에 대해 Handcraft feature 또는 SSD [8]로 추가적인 검출과 분류를 수행했다. Zeng 등 [9] 역시 다중 센서를 기반으로 사전 지도를 생성하고 GPS로 위치 인식을 하여 신호등 위치를 추정한 뒤, 최종적으로 YOLOv4를 사용하여 신호등을 인식하였다. 이 방법들은 사전정보 기반으로 신호등 영역을 찾은 후에 객체 검출을 진행하므로 추가 비용이 들고, 다른 연구와 성능 비교를 하지 않았다.

본 논문에서는 Hirabayashi 등 [6]의 연구와 유사하게 Autoware 프레임워크를 기반으로 한 다단계 접근법을 제안한다. 기존 연구와 달리, 분류 성능과 속도 개선을 위해 검출 기능이 없는 ResNet-50 [10]와 PANet [11] 기반의 분류기 모델을 제안하며, 모델의 공정한 평가를 위하여 공개된 신호등 데이터 셋인 LISA [12]를 사용하여 정량평가로 성능 비교를 수행하였다.

본 논문의 기여는 다음과 같다. 타겟 시나리오에 적합한 센서 구성 및 설치 방법을 제안한다. 보다 효과적이고 균형된 데이터 셋 확보를 위한 자동화된 데이터 셋 구축 프로세

스를 제안한다. 또한 자율주행의 필수 요소 (위치 인식, 맵)를 기반으로 신호등 분류 모델만 추가하여 속도와 인식 성능을 향상시키는 방법을 제안하고, 공개 데이터 셋 기반의 정량평가와 성능 비교를 제시한다.

III. 신호등 인식 모듈 설계

1. 타겟 시나리오 기반의 센서 환경구축

신호등 체계는 국가별, 지역별 특성에 따라 그 크기와 종류 (종/횡형, 구 개수 등)가 다양하며, 도로 환경에 따라 차선의 수와 폭, 최고 시속 등이 다르게 나타난다 [13]. 따라서 타겟 시나리오를 선정하는 것은 매우 중요하다. 시나리오 구간을 분석하여 식별해야 할 신호등의 종류와 인식 거리, Field Of View (FOV) 등 검출 범위를 결정해야 한다.

본 논문에서는 그림 1에서와 같이 대구 소도심 지역인 테크노폴리스 내의 약 4.5km, 편도 3~4차로가 있는 도로를 시나리오 구간으로 선정하였다. 타겟 시나리오 분석에 따른 요구 사양은 표 1과 같다. 요구 사양 중에서 신호등 수, 종류, 최고 시속, $traffic_h$, $traffic_{dist}$, $lane_{width}$ 는 시나리오 구간의 주어진 환경에 따르며, cam_h 와 $bumper_{dist}$ 는 카메라가 설치된 위치에 의해 결정된다. $range_{max}$ 는 최고 시속 시 제동거리와 반응거리 (반응시간 0.1s 기준), 그리고 정지선과 교차로 너머 설치된 제2 신호등 간의 최대 거리를 고려하여 70m로 설정하였다.

이미지에서 신호등을 식별할 수 있는 최소 픽셀 크기 (px_{min})는 10픽셀로 정하였고 수식 (1)을 이용하여 신호등 램프 ($lamp_size_m$)가 70m ($range_{max}$)에 있을 때, 이미지에 10px (px_{min}) 이상 만족하는 이미지 센서 픽셀 크기 (px_{size})



그림 1. 타겟 시나리오 영역
Fig. 1. Target scenario area

표 1. 타겟 시나리오의 요구 사양

Table 1. Requirements for the target scenario

항목	수치	기준
Weather	-	sunny and daytime
# of TL	14 Set	TLs with the same signal are counted as one
Type	6	G, R, Y, Left, Straight-Left, Unknown
$lamp_size_m$	0.355 m	TL lamp size according to regulations
$traffic_h$	6.3 m	Height of the highest TL
$traffic_{dist}$	3.5 m	Min. distance between stop line and TL
$lane_width$	3.3 m	Maximum lane width
cam_h	1.7 m	Camera height installed in the vehicle
$bumper_{dist}$	2.3 m	Distance between bumper and camera
$range_{max}$	70 m	The farthest distance between TL and stop line(50 m) + braking distance(18 m) + reaction distance(2 m)
Max. speed	60 km/h	Max. speed according to regulations
VFOV	38.4° Eq. (2)	Tilt between TL and FOV of the camera on stop line
HFOV	59.2° Eq. (2)	Yaw between TL and FOV of the camera on stop line

표 2. 선정된 카메라 모델

Table 2. Selected camera model

Model name	BFS-PGE-50S5C-C
Resolution (WxH)	2448 x 2048
Sensor pixel size (px_size)	3.45 $\mu\text{m}/\text{px}$
FPS	10 frame/sec
Lens focal length (f_{mm})	8 mm
VFOV	47.6°
HFOV	55.6°

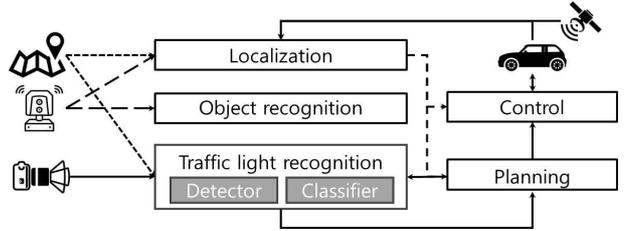


그림 3. 자율주행 플랫폼 아키텍처
Fig. 3. Architecture of autonomous platform

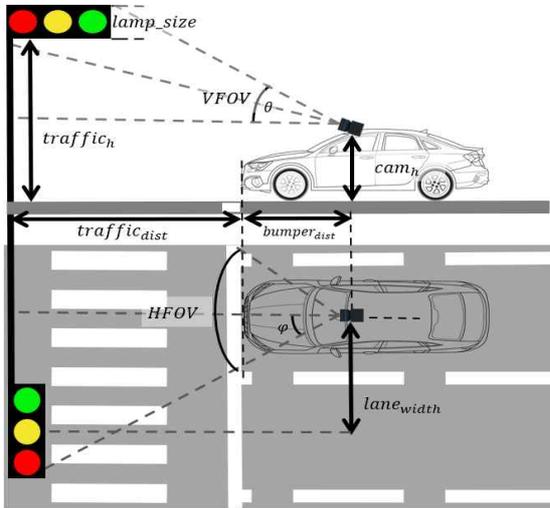


그림 2. 카메라 FOV의 기하학 모델
Fig. 2. Geometric model of camera FOV

와 렌즈의 focal length (f_{mm}) 조합을 계산할 수 있다. VFOV와 HFOV는 수식 (2) 을 통해 정지선에 차량이 정차했을 때 FOV에 근접한 제1 신호등을 포함하는 최대 tilt (θ) 와 yaw (φ)로 결정된다 (그림 2 참조). 표 1의 기준을 만족하는 카메라 모델을 표 2와 같이 선정하였으며, 요구 사양 대비 부족한 카메라의 HFOV는 비교적 원거리에 있는 제2 신호등만 인식하는 것으로 설계하였다.

$$f_{mm} = px_size \cdot px_min \cdot \frac{range_{max}}{lamp_size_m}, \quad (1)$$

$$VFOV = \arctan\left(\frac{traffic_h - cam_h}{traffic_{dist} + bumper_{dist}}\right), \quad (2)$$

$$HFOV = 2 \cdot \arctan\left(\frac{lane_width}{traffic_{dist} + bumper_{dist}}\right).$$

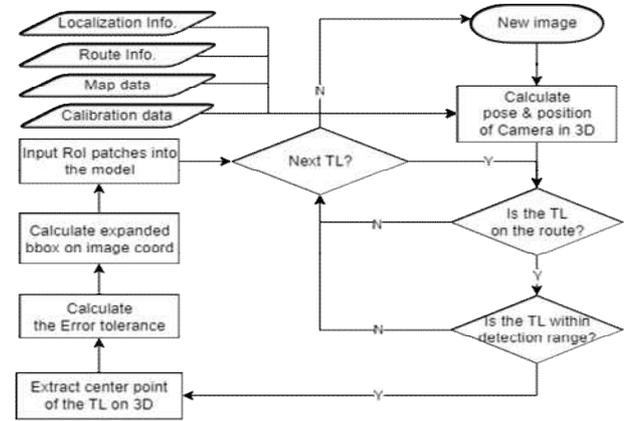


그림 4. 신호등 검출 흐름도
Fig. 4. TL detection flow

2. 아키텍처 설계

본 논문에서 제안하는 신호등 인식 모듈은 자율주행을 위한 open source software project인 Autoware [7] 기반에서 동작하며, 그림 3과 같이 위치와 맵 정보를 기반으로 이미지에서 신호등 RoI를 찾는 Detector와 이를 분류하는 Classifier로 구성된다. 이 아키텍처를 통해 신호등 인식 모듈이 단독으로 극복할 수 없는 예외 상황 (자차 신호등 결정, 미인식된 신호등에 대한 대처와 딜레마 존에서의 대응 등)에 대해서는 단순한 규칙 기반 결정으로 해결할 수 있으며, 이미지 RoI 기반으로 신호등을 인식하므로 처리 속도와 인식 성능을 높일 수 있다. 인식 효율을 극대화하기 위해 그림 4의 전처리 과정을 거치며, 주행 상황에서 발생하는 localization, calibration, synchronization, vibration 등의 오차를 극복하도록 허용 오차 (Error tolerance)로 확장된 RoI를 생성하는 기법을 적용한다. 수식 (3)과 같이 검출된 맵 좌표계의 신호등 중심점을 카메라 좌표계로 변환한 뒤, 이

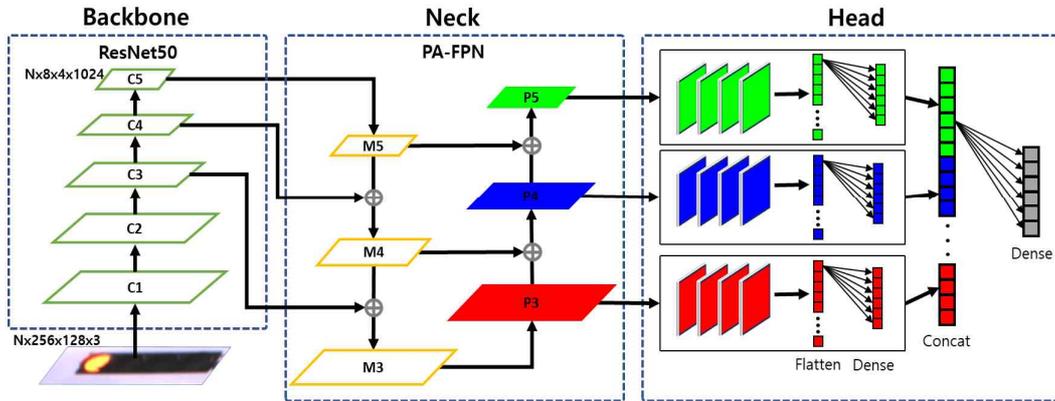


그림 5. 제안하는 신호등 분류 모델의 구조

Fig. 5. Structure of the proposed traffic light classification model

중심점에 허용 오차를 적용하여 좌상단, 우하단 점을 계산하고 이미지에 투영하여 RoI를 계산한다 (“±”가 -인 경우 좌상단, +인 경우 우하단). 수식에서 et_ϕ , et_θ , et_x , et_y , et_z 는 허용 오차이며, 실험에서 et_ϕ 와 et_θ 는 각각 0.5도로, et_x , et_y , et_z 는 각각 0.05m로 지정하였다.

$$\begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X_{map} \\ Y_{map} \\ Z_{map} \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} X_{cam}^{expand} \\ Y_{cam}^{expand} \\ Z_{cam}^{expand} \end{bmatrix} = \begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \end{bmatrix} \pm \left(Z_{cam} \begin{bmatrix} \sin(et_\phi) \\ \sin(et_\theta) \\ 0 \end{bmatrix} + \begin{bmatrix} et_x \\ et_y \\ et_z \end{bmatrix} \right), \quad (3)$$

$$\begin{bmatrix} RoI_x \\ RoI_y \\ 1 \end{bmatrix} = \frac{1}{Z_{cam}^{expand}} \cdot \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_{cam}^{expand} \\ Y_{cam}^{expand} \\ Z_{cam}^{expand} \end{bmatrix}.$$

3. 모델 디자인

제안하는 신호등 분류 모델은 Backbone, Neck, Head로 구성된다 (그림 5 참조). Backbone은 ResNet-50 [10]을, Neck은 PANet [11]을 채택하여 low-level과 high-level 특징 간의 정보가 잘 전달될 수 있도록 설계하였다. Head는 신호등 6종을 분류하는 태스크로서, 기존의 검출기와 달리 Flatten과 Dense layer만으로 구성하였다. 신호등의 특성상 거리에 따라 어떤 feature pyramid에서 유용한 형상과 색상 정보가 배치될지 알 수 없으므로, multi-scale feature pyramid를 하나로 통합하지 않고 각 level에서 flatten과 dense layer를 통해 균등한 feature output을 얻는다. 이후, concat과 dense layer를 통해 최종 결과를 얻는다.

손실함수는 클래스 불균형 문제를 고려하여 focal loss [14]를 사용하였으며, 클래스는 표 1의 종류와 같다. 입력 이미지 크기 (256x128)는 3.2절에서 설명한 확장된 RoI의 평균 중형비를 취득한 데이터 셋에서 찾고 (2:1), 이를 유지하면서 최종 convolution의 특징이 소멸하지 않는 크기 (8x4)를 갖도록 입력 크기를 역으로 계산하여 결정하였다.

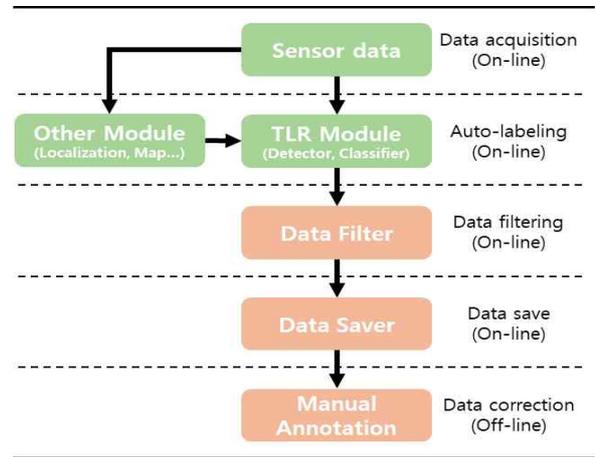


그림 6. 제안하는 데이터 셋 구축 프로세스

Fig. 6. Proposed dataset construction process

IV. 신호등 데이터 셋 구축 프로세스 설계

1. 아키텍처 설계

신호등 데이터 셋 구축 프로세스는 신호등 분류 모델의 학습과 평가를 위한 데이터 셋 구축 과정을 자동화한 온라인 프로세스이다 (그림 6 참조). 그림 3의 아키텍처를 그대로 활용하여 센서 데이터를 취득하고, 신호등 인식 모듈로 자동 라벨링을 수행한 뒤, 다음 절에서 설명할 데이터 필터링과 저장, 수동 보정 작업을 거쳐 최종적으로 데이터 셋을 자동으로 생성한다. 데이터 취득, 라벨링, 필터링 및 저장 과정은 자율주행 상태에서 온라인으로 수행되며, 데이터 보정 작업만 오프라인에서 수행된다 [15].

2. 데이터 필터 및 저장

데이터 필터는 분류 모델의 관점에서 “어려운 샘플”과 “다양한 샘플”을 선별하는 과정이다. 각 샘플의 정의는 다음과 같다.

- 어려운 샘플 : 모델이 낮은 신뢰도 값 (Confidence)을 출력하는 샘플

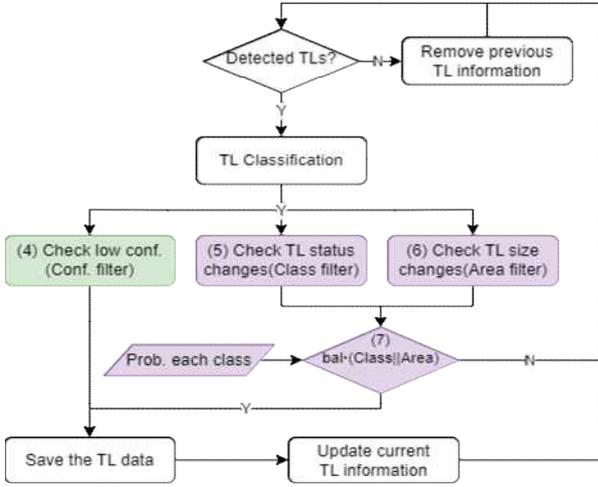


그림 7. 데이터 필터링과 저장 흐름도 (괄호의 숫자는 수식 번호)
Fig. 7. Data filtering and saving flow (The number is a formula number)

• 다양한 샘플 : 직전 프레임의 신호등 결과와 비교했을 때 신호가 변경되거나, 이전 비교 시점 대비 신호등 RoI가 크게 변한 샘플 중에서, 클래스 균형을 고려하여 선별된 샘플
그림 7은 데이터 필터링과 데이터 저장 과정의 흐름도를 나타내며, 신호등 인식 결과가 모델로부터 추론되면 데이터 필터링을 위한 각 필터를 통과한 데이터만 저장하는 과정을 보여준다. “어려운 샘플”의 추출 과정은 $Filter_{conf}$ 으로 판단되며, 수식 (4)과 같이 모델이 출력한 신호등의 신뢰도 값이 특정 임계값 (α)이하인 경우로 한정된다.

“다양한 샘플”의 추출 과정은 두 단계로 구성된다. 먼저 1 단계에서는 수식 (5)의 $Filter_{class}$ 을 통해 직전 프레임에서 인식한 신호등 결과와 비교하여 신호 변경 여부를 감지하며, 수식 (6)의 $Filter_{area}$ 을 통해 이전 비교 시점 프레임의 신호등 RoI 영역과 현재 프레임의 RoI 영역의 변화가 특정 임계값 (β)이상인 경우를 감지한다. 이 두 조건 중 하나라도 참일 때 2단계로 진행된다. 2단계에서는 수식 (7)의 $Filter_{balance}$ 을 통해 클래스 균형을 고려하기 위해 $Prob_{class}$ 을 적용하여 데이터의 저장 여부를 최종결정한다.

$$Filter_{conf} = \begin{cases} 1 & \text{if } confidence_i^{ID} \leq \alpha, \\ 0 & \text{else} \end{cases} \quad (4)$$

$$Filter_{class} = \begin{cases} 1 & \text{if } class_i^{ID} \neq class_{i-1}^{ID}, \\ 0 & \text{else} \end{cases} \quad (5)$$

$$Filter_{area} = \begin{cases} 1 & \text{if } |area_i^{ID} - area_{i-m}^{ID}| \geq \beta, \\ 0 & \text{else} \end{cases} \quad (6)$$

$$Filter_{balance} = Prob_{class} \cdot (Filter_{class} \mid \mid Filter_{area}), \quad (7)$$

여기서, $Prob_{class} = \frac{count_{max} - count_{class}}{count_{max} - count_{min}}$.

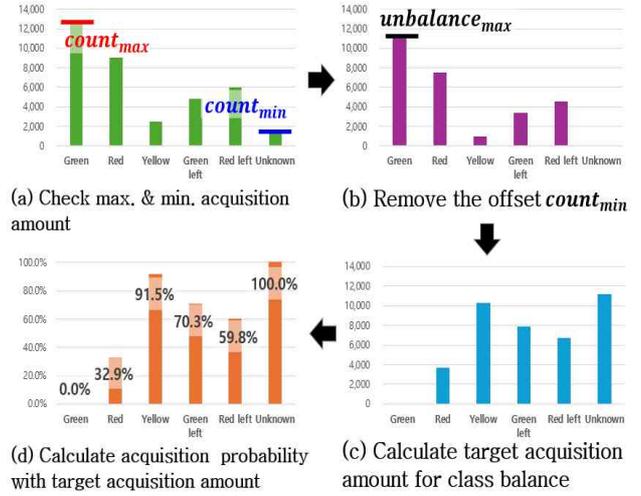


그림 8. 클래스별 취득 확률 결정 과정
Fig. 8. Acquisition probability determination process for each class

표 3. 각 필터의 파라미터
Table 3. Parameters of each filter

Confidence threshold (α)		90.0%
Area threshold (β)		700 px ²
$Prob_{class}$	Green	0.0%
	Red	32.9%
	Yellow	91.5%
	Green Left	70.3%
	Red Left	59.8%
	Unknown	100.0%

수식 (7)의 $Prob_{class}$ 은 그림 8의 과정을 거쳐 도출된 클래스별 취득 확률을 의미한다. 먼저 기취득된 데이터 셋 분포를 분석하여 최대 ($count_{max}$)와 최소 ($count_{min}$)로 수집된 클래스 양을 확인한다 (그림 8 (a)). 클래스별로 불균형 크기를 계산하기 위해 오프셋인 $count_{min}$ 을 각 클래스의 취득한 양에서 제거한다 (그림 8 (b)). 클래스 균형을 달성하기 위해 그림 8(b)의 $unbalance_{max}$ 으로 각 클래스의 목표 취득 양을 계산한다 (그림 8 (c)). 마지막으로, 각 클래스의 목표 취득한 양과 $unbalance_{max}$ 간 비율 이용해 취득 확률을 계산한다 (그림 8 (d)). 이는 취득된 신호등 데이터를 클래스별로 저장 여부를 결정하는 확률값으로 사용된다. 실험에 사용한 필터 파라미터는 표 3과 같다.

3. 데이터 보정

데이터 보정은 자동 라벨링 과정에서 가공 오류가 있는 데이터들을 수정하는 작업으로써, 이는 오프라인에서 수행된다. 자체 제작한 툴로 수동 보정 작업을 마친 데이터는 신호등 분류 모델의 학습과 평가를 위한 데이터 셋으로 활용된다.

V. 실험 결과

1. 타겟 시나리오 평가를 위한 데이터 셋 구축

표 4. 자체 데이터 셋 구축량

Table 4. The amount of our dataset

Data type	Total	Train		Test	
	count	count	ratio	count	ratio
Image	18,033	15,275	-	2,758	-
Green	12,695	10,764	34.7%	1,931	34.3%
Red	9,022	7,658	24.7%	1,364	24.2%
Yellow	2,491	2,129	6.9%	362	6.4%
Green Left	4,848	4,206	13.6%	642	11.4%
Red Left	6,028	5,018	16.2%	1,010	17.9%
Unknown	1,538	1,218	3.9%	320	5.7%

표 5. 자체 데이터 셋에 대한 정량적 평가 결과

Table 5. Quantitative results of our dataset

Metric	AP (%)	F1 (%)	AUC (%)
Green	99%	99%	99.8%
Red	99%	100%	100%
Yellow	98%	99%	98.7%
Green Left	99%	99%	99.5%
Red Left	99%	100%	100%
Unknown	92%	92%	94.1%
Average	98%	98%	98.6%

신호등 체계는 국가별, 지역별 특성에 따라 종류가 다양하다. 따라서 이 논문에서는 타겟 시나리오에 강건한 인식 모듈을 개발 및 평가하기 위해 그림 1의 실제 환경에서 표 1의 요구사항을 반영하여 데이터를 취득하였다. 또한 3장에서 설명한 데이터 셋 구축 프로세스인 데이터 취득, 필터, 보정 과정을 거쳐 표 4에서와 같이 자체 데이터 셋을 구축하였으며, 평가 데이터 셋은 학습 데이터 셋과 유사한 비율을 가지도록 구성하였다.

2. 자체 데이터 셋 기반 평가

모델의 학습은 8대의 GPU를 사용하며, GPU 당 배치 크기는 128, 최대 반복은 60k, 초기 학습률은 0.03에서 반복이 40k와 50k를 달성될 때 각 0.003, 3.0e-4로 스텝 다운을 수행하였다. 평가는 AP (Average Precision)와 F1 score (수식 8)를 이용한다 [2, 4, 9]. 정량평가 결과는 표 5와 같다 (그림 9 참조).

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (8)$$

3. 공개 데이터 셋 기반 평가

본 연구에서는 제안한 모델의 객관적 정량평가를 위해 잘 알려진 신호등 공개 데이터 셋인 LISA [12] 활용하였다. LISA를 포함한 여러 신호등 공개 데이터 셋은 맵 정보를 제공하지 않으므로, 이를 대신하여 [9]에서와 같이 생성한 RoI 기반 평가를 수행하였다. RoI는 3.2절에서 설명한 확장된 RoI와 정확한 신호등 RoI 간의 관계를 자체 데이터 셋에서 width와 height, 중심점을 각각 모델링하고 이를 LISA 데이터 셋에서 확장된 RoI로 생성하여 정량평가를 수행하였다.

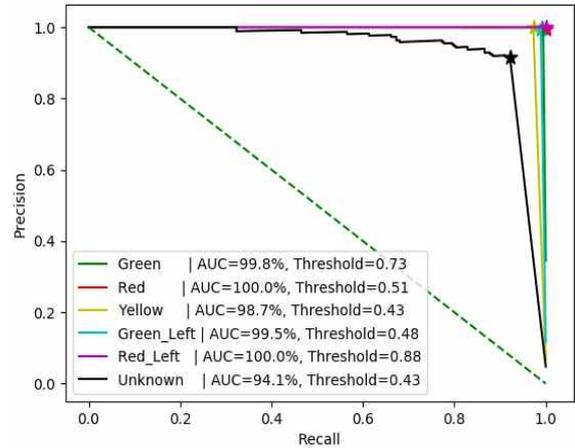


그림 9. Precision-Recall 곡선
Fig. 9. Precision-Recall Curve

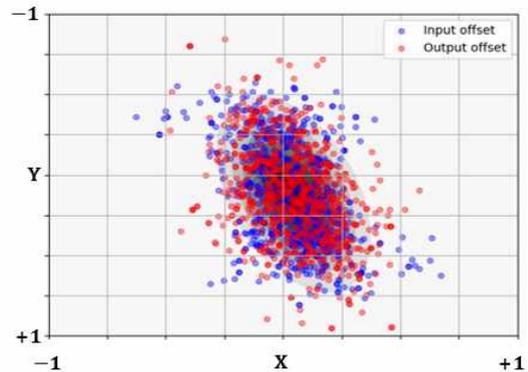


그림 10. 생성한 오프셋 분포
Fig. 10. Generated offset distribution

3.1 확장된 RoI 모델링

확장된 RoI의 width와 height는 정확한 신호등 RoI의 width와 height를 입력으로 하는 선형회귀 (Linear regression) 방법을 통해 1차 방정식으로 각각 모델링한다. 중심점의 모델링은 수식 (9)과 같이 확장된 RoI와 신호등 RoI의 중심점 간의 오프셋을 확장된 RoI의 width와 height로 정규화하여 -1부터 +1 사이의 값을 가지도록 모델링하였다 (그림 10 참조). 그리고 이 오프셋은 수식 (10)의 다변량 정규분포 확률로 모델링한다.

$$center^{off} = center^{expand} - center^{TL}, \quad (9)$$

$$center^{norm} = \frac{center^{off}}{0.5 \cdot wh^{expand}}$$

$$f(x) = \frac{1}{\sqrt{(2\pi)^2 \det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right). \quad (10)$$

3.2 확장된 RoI 생성

LISA 데이터 셋은 세로형 신호등으로 구성되어 있으므로, x, y 좌표 및 width, height 값을 반대로 적용하여 생성한다. 확장된 RoI의 생성 순서는 다음과 같다. 먼저 LISA 데이터

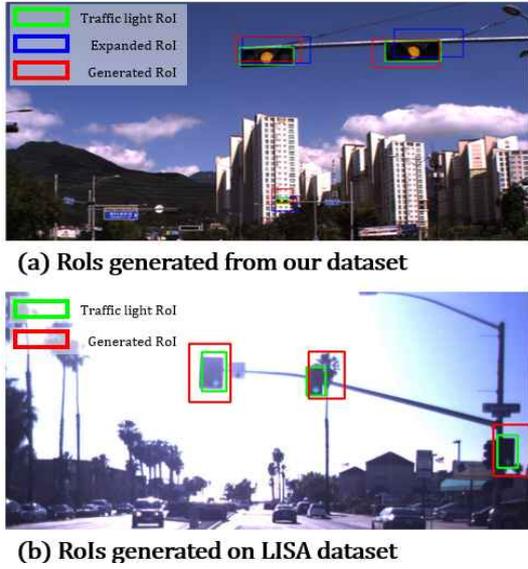


그림 11. 생성된 RoI 예시
Fig. 11. Generated Rols example

셋의 신호등 RoI width와 height를 선형회귀 모델에 입력하여 확장된 RoI의 width와 height를 각각 생성한다. 그 후, 수식 (10)을 이용하여 정규분포를 따르는 중심점 오프셋을 생성한다. 이를 수식 (11)을 통해 확장된 RoI의 중심점으로 변환한다. 이렇게 구해진 width, height, 중심점을 통해 그림 11과 같은 확장된 RoI를 생성할 수 있다.

$$\begin{aligned}
 center_{gen}^{off} &= 0.5 \cdot wh_{gen}^{expand} \cdot center_{gen}^{norm}, \\
 center_{gen}^{expand} &= center^{TL} + center_{gen}^{off}.
 \end{aligned}
 \tag{11}$$

3.3 정량평가

LISA의 학습 및 평가 데이터 셋은 [4]와 동일하게 나누어 평가를 수행하였다. 모델은 8대의 GPU를 사용하였으며, GPU 당 배치 크기는 32, 최대 반복 횟수는 20k, 초기 학습률 4.0e-4에서 학습이 16k와 18k에 도달할 때 각각 4.0e-5, 4.0e-6으로 스텝 다운하여 학습하였다. 평가 결과는 표 6에 제시되었으며, 기존 연구와 비교하여 속도와 정확도의 모든 성능 지표에서 개선된 것을 확인하였다.

표 6. LISA 데이터 셋에 대한 정량적 평가 결과
Table 6. Quantitative evaluation results of LISA dataset

* Ours : 57개 RoI를 batch 처리, NVIDIA Titan Xp
* [4] : NVIDIA GTX 1080Ti

Method	AP (%)			mAP (%)	ms*
	Go	Stop	Warning		
Ours	99.5	99.6	93.9	97.6	19.2
[4]	76.6	91.2	78.2	82.1	33.7

VI. 결론

다양한 신호등이 있는 국내 환경에서 도심 자율주행을 위한 신호등 인식 기술은 매우 중요하다. 우리는 자율주행에 필수 불가결 요소들 (맵, 센서)을 이용하여 타겟 시나리오 기반의 고성능 신호등 인식 모듈을 설계하였다. 이 논문에서는 카메라의 의존성을 최소화하기 위해, 위치 인식 기술과 맵을 활용하여 신호등의 정확한 위치 정보를 기반으로 카메라 이미지에서 신호등 RoI를 추출해 인식 성능은 높이고 처리 시간은 줄이는 데 주력하였다. 저전력과 분산처리를 위해 차량용 PC와 Jetson xavier 보드를 활용하였고, 차량용 PC에서는 신호등 인식 모듈 외에 위치 인식과 맵을 포함하는 모듈이 구동되며, 이 모듈들이 위치 정보와 맵의 신호등 정보를 기반으로 신호등 RoI를 추출해 신호등 분류 모듈이 있는 Xavier에 전송한다. Xavier에서 추론 결과와 다시 차량용 PC에 전달하여 자율주행을 수행하도록 구성하였다. Xavier에서 신호등 1개의 RoI를 처리하는 시간은 평균 43.2ms 소요되지만, 타겟 시나리오상에서는 한 장면에서 자차 신호등이 최대 3개 나타날 수 있으므로 이를 batch 추론을 통해 10 fps 이내 처리 시간 (평균 93.4 ms)을 달성하였다. 또한, 공개 데이터 셋 기반의 정량평가로 기존 연구와 성능 비교를 통해 모델의 우월성을 검증하였고, 임베디드 보드에서 자율주행 실증을 통해 효율성도 입증하였다.

효율적인 적용을 위해 자동화된 데이터 셋 구축 프로세스도 제안하였다. 데이터 취득, 자동 라벨링, 데이터 필터링 및 저장 과정은 자율주행 상태에서 온라인으로 수행되며, 데이터 보정 작업만 오프라인에서 수행된다. 이를 통해 낮은 비용으로 클래스 균형을 유지하는 다량의 데이터 셋을 손쉽게 구축하여 모델의 성능을 높일 수 있다.

References

- [1] C. Jung, D. Lee, S. Lee, D. H. Shim, "V2X-communication-aided Autonomous Driving: System Design and Experimental Validation," *Sensors* Vol. 20, No. 10, pp. 2903, 2020.
- [2] M. Weber, P. Wolf, J. M. Zöllner, "DeepTLR: A Single Deep Convolutional Network for Detection and Classification of Traffic Lights," 2016 IEEE Intelligent Vehicles Symposium (IV). IEEE, pp. 342-348, 2016.
- [3] J. Redmon, S. Divvala, R. Girshick, "You Only Look Once: Unified, Real-time Object Detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 779-788, 2016.
- [4] Q. Wang, Q. Zhang, X. Liang, Y. Wang, C. Zhou, V. I. Mikulovich, "Traffic Lights Detection and Recognition Method Based on the Improved YOLOv4 Algorithm," *Sensors*, Vol. 22, No. 1, pp. 200, 2021.
- [5] M. Mostafa, M. Ghantous, "A Yolo Based Approach for Traffic Light Recognition for Adas Systems," 2022 2nd

- International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC). IEEE, pp. 225-229, 2022.
- [6] M. Hirabayashi, A. Sujiwo, A. Monroy, S. Kato, M. Edahiro, "Traffic Light Recognition Using High-definition Map Features," *Robotics and Autonomous Systems* 111, pp. 62-72, 2019.
- [7] S. Kato, E. Takeuchi, Y. Ishiguro, Y. Ninomiya, K. Takeda, T. Hamada, "An Open Approach to Autonomous Vehicles," *IEEE Micro*, Vol. 35, No. 6, pp. 60-68, 2015.
- [8] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, A. C. Berg, "Ssd: Single Shot Multibox Detector," *Computer Vision - ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 - 14, 2016, Proceedings, Part I 14*. Springer International Publishing, pp. 21-37, 2016.
- [9] Z. Li, Q. Zeng, Y. Liu, J. Liu, L. Li. "An Improved Traffic Lights Recognition Algorithm for Autonomous Driving in Complex Scenarios," *International Journal of Distributed Sensor Networks*, Vol. 17, No. 5, 15501477211018374, 2021.
- [10] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for ImageRecognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Vol. 34. pp. 770-778, 2016.
- [11] S. Liu, L. Qi, H. Qin, J. Shi, J. Jia, "Path Aggregation Network for Instance Segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 8759-8768, 2018.
- [12] M. B. Jensen, M. P. Philipsen, A. Møgelmoose, T. B. Moeslund, M. M. Trivedi, "Vision for Looking at Traffic Lights: Issues, Survey, and Perspectives," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 7, pp. 1800-1815, 2016
- [13] Korean National Police Agency, "Traffic Signal Installation and Operation Manual," 2022.
- [14] T. Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, "Focal Loss for Dense Object Detection," *Proceedings of the IEEE International Conference on Computer Vision*. pp. 2980-2988, 2017.
- [15] J. H. Lee, J. K. Lee, J. Park, J. S. Kim, S. Kwon, "Construction and Effectiveness Evaluation of Multi Camera Dataset Specialized for Autonomous Driving in Domestic Road Environment," *IEMEK J. Embed. Sys. Appl.*, Vol. 17, No. 5, pp. 273-280, 2022.

Jaehyeong Park (박재형)



2013 Electronic control engineering from Daegu University of Korea (B.S.)
 2015 Electronic Control and Measurement from Daegu University (M.S.)
 2015~Division of Automotive Technology from Daegu Gyeongbuk Institute of Science & Technology (DGIST). (Researcher)

Field of Interest: Autonomous driving, Deep learning, Sensor fusion
 Email: stillrunning@dgist.ac.kr

Jin-Hee Lee (이진희)



2005 Computer Science from Korea National Open University (B.S.)
 2007 Computer and Information Engineering from Inha University (M.S.)
 2015 Computer and Information Engineering from Inha University (Ph.D.)

2015~Division of Automotive Technology from Daegu Gyeongbuk Institute of Science & Technology (DGIST). (Senior researcher)

Field of Interest: Autonomous driving, Deep learning, Sensor fusion
 Email: jhlee07@dgist.ac.kr

Je-Seok Kim (김세석)



2009 Mechanical Engineering from Yeungnam University (B.S.)
 2012 Automotive Engineering from Hanyang University (M.S.)
 2015 Automotive Engineering from Hanyang University (Ph.D.)

2020~Division of Automotive Technology from Daegu Gyeongbuk Institute of Science & Technology (DGIST). (Senior researcher)

Field of Interest: Automotive driving, Planning algorithm
 Email: jeseok@dgist.ac.kr

Soon Kwon (권순)



2003 Electric & Electronic Engineering from Korea University (B.S.)
 2006 Electronic Engineering from Seoul National University (M.S.)

2006~Division of Automotive Technology from Daegu Gyeongbuk Institute of Science & Technology (DGIST). (Principal researcher)

Field of Interest: Autonomous driving, Deep learning, Sensor fusion
 Email: soonyk@dgist.ac.kr