

CCTV 영상을 활용한 동적 객체의 위치 추적 및 시각화 방안

Location Tracking and Visualization of Dynamic Objects using CCTV Images

박상진* · 조국** · 임준혁*** · 김민찬****
Park, Sang-Jin · Cho, Kuk · Im, Junhyuck · Kim, Minchan

Abstract

C-ITS(Cooperative Intelligent Transport System) that pursues traffic safety and convenience uses various sensors to generate traffic information. Therefore, it is necessary to improve the sensor-related technology to increase the efficiency and reliability of the traffic information. Recently, the role of CCTV in collecting video information has become more important due to advances in AI(Artificial Intelligence) technology. In this study, we propose to identify and track dynamic objects(vehicles, people, etc.) in CCTV images, and to analyze and provide information about them in various environments. To this end, we conducted identification and tracking of dynamic objects using the Yolov4 and Deepsort algorithms, establishment of real-time multi-user support servers based on Kafka, defining transformation matrices between images and spatial coordinate systems, and map-based dynamic object visualization. In addition, a positional consistency evaluation was performed to confirm its usefulness. Through the proposed scheme, we confirmed that CCTVs can serve as important sensors to provide relevant information by analyzing road conditions in real time in terms of road infrastructure beyond a simple monitoring role.

Keywords: Autonomous Driving, Road Infrastructure, CCTV, AI, Coordinate System Transformation.

1. 서론

자율주행차의 상용화가 이동성과 편의성 측면에서 긍정적인 영향을 미칠 것으로 기대됨에 따라 최근 차량의 센서정보를 활용한 제어 알고리즘 및 관련 기술을 중심으로 자동차 제조사, IT 업계에서 많은 연구개발이 진행되고 있다. 그러나, 이러한 차량의 센서만을

로 주변 환경을 인지하기에는 센서 사각지대, 악천후 등 안전성 측면에서 한계가 존재한다. 따라서 안전한 자율주행 기술의 구현을 위한 효율적이고 유용한 도로 인프라가 필요하다. 자율주행 인프라는 통신을 통해 자율주행 차량에게 날씨, 공사정보, 주변 차량, 보행자 정보를 제공하며, 정밀측위를 지원하는 등 자율주행 차량의 주행을 지원한다. 도로 인프라는 Table 1

* 공간정보연구원 선임연구원 LX Spatial Information Research Institute (first author: sangjinpark@lx.or.kr)

** 공간정보연구원 선임연구원 LX Spatial Information Research Institute (co-author: kcho@lx.or.kr)

*** 공간정보연구원 선임연구원 LX Spatial Information Research Institute (co-author: junhyuck@lx.or.kr)

**** 공간정보연구원 선임연구원 LX Spatial Information Research Institute (corresponding author: minchankim@lx.or.kr)

Table 1. Classification of Road Infrastructure.
Source: Park et al. 2019

	기술분류	요소기술
인프라	도로시설물	스마트 톨게이트, 스마트 신호등, 발광 차선 등의 자율주행 지원 도로시설물
	노변센서	보행자, 차량, 장애물 등을 감지하는 카메라, 레이더 등의 센서
	교통센터	교통신호, 공사, 사고 등의 정보 수집·관리
	통신	4/5G, WAVE 등
	기타	기획/전략/보험 등 연구

에서와 같이 도로시설물, 노변센서, 교통센터, 통신, 기타로 분류할 수 있다.

도로시설물은 자율주행의 인지성능 향상과 사고위험 감소 등을 위해서 도로에 적용된 시설물을 의미하며, 노변센서는 다양한 센서를 통해 도로 주변의 물체와 환경을 감지하는 기술을 의미한다. 교통센터는 차량·도로시설물·노변센서 등에서 수집된 데이터를 종합적으로 관리·분석하는 기술을 의미하고, 자율주행을 위해 수집·분석된 데이터를 차량-차량 간, 차량-인프라 간에 송수신하는 통신기술 또한 도로 인프라에 속하는 중요한 기술이다. 이외에 도로 인프라에는 구축을 위한 기획/전략 등을 연구하는 분야도 포함하고 있다.

이러한 도로 인프라 기술의 향상을 위해 최근 차세대 지능형 교통 체계(C-ITS: Cooperative -ITS) 등과 같은 자율주행 인프라 관련 사업들이 수행되고 그 안에서 많은 연구개발이 진행되고 있으며, 그 결과물들을 활용한 실증을 수행하고 있는 단계에 있다.

C-ITS란 기존 지능형 교통 체계(ITS: Intelligent Transport Systems)의 한계를 극복하기 위해 기존 시스템에 차량간 통신(V2X) 를 포함하여 교통안전에 더욱 중점을 둔 시스템을 의미한다(Kiela et al. 2020).

ITS는 센터 중심(차량-센터)의 단방향 교통정보 제공 서비스로 차량-차량, 차량-센터 간정보 수집과 제공을 분리해야 되는 한계가 존재하는 반면, C-ITS는 주행(차량-차량, 차량-센터) 중심의 양방향 교통정보 제공이 가능하며 주변 차량 및 도로 상황과 연속적인 상호 통신을 통해 교통정보를 교환 및 공유할 수 있다. 이는 사고, 공사 등과 같은 교통상황 변화에 따른 주행 중심의 신속한 대응이 가능하다는 이점을 가지며 자율주행의 한계를 극복하기 위한 도로 인프라로 규정되고 있다. 이러한 C-ITS 기술은 차량에 탑재된 센서에만 의존했던 부분을 인프라와 주변차량, 동적객체(보행자 등)들의 정보를 이용하여 보다 효율적이고 신뢰성 있는 자율주행을 지원할 수 있다는 평가가 주를 이루고 있어 국외(Bezzina et al. 2015; CV Pilot 2017; Kotsi A 2020; ACM 2017)뿐만 아니라 국내(김진태 2018)에서도 연구개발과 시범사업을 활발하게 추진하고 있는 실정이다.

C-ITS의 도로 인프라 센서에서 물체와 상황 등을 인지할 수 있는 카메라, 레이더, 라이다 등은 자율주행 차에도 적극적으로 도입될 만큼 자율주행에 있어 활용도가 높은 센서에 속한다. 그중 카메라 영상 센서에 해당하는 CCTV의 경우, 도로 상태 및 상황, 보안 등을 관리하는 센터에서 모니터링의 필요성에 의해 구간별로 많은 양을 구축하여 운영하고 있으나, 활용도가 다양함에도 불구하고 단순한 영상 모니터링에 주로 활용되고 있어 자율주행 측면에서 활용도는 타 센서들에 비해 부족한 실정이다. CCTV를 통해 차량, 보행자 등 이동체를 인식하고 지도상 위치 정보를 도출 한다면, 이는 사각지대 돌발 상황 검출, 보행자 케어 등 다양한 C-ITS 서비스에 다양한 형태로 접목·활용할 수 있을 것이다. 이와 관련하여 도로교통 정보 분석을 위한 딥러닝 기반 객체 인지 및 영상분석에 대한 연구는 꾸준히 진행되고 있으며, 현재까지는 주로 차량, 사람 등 특정 객체의 인식률을 향상시키거나, 움직임 분석하기 위한 연구들이 진행되고 있어 인지된 객체의

좌표계 변환을 통해 새로운 형태로의 시각화 등 활용·응용을 위한 연구는 아직 미비한 실정이다.

양인철 등(2019)과 서흥덕 등(2020)은 대용량 데이터 처리를 위해 구글 클라우드 플랫폼을 활용하여 영상에서 차량들을 인지하거나, 스테레오 CCTV를 통해 인지된 차량의 위치정보를 추정하여 해당지역의 교통량을 추정하는 방안을 제안하였으며, Mehboob 등(2017)은 CCTV영상에서 차량을 추적하고 변환관계를 통해 추적결과를 구글맵과 연계하여 이벤트 정보 맵에 시각화 하여 관련 정보를 제공하는 방안을 제안하였다. 또한, Sankaranarayanan 등(2008)은 스테리오 비전 방식을 채용한 6대의 카메라 간 변환관계를 통해 특정 지역에서의 객체 추적 및 위치 변환을 통한 탐부에서의 객체 시각화 방안을 제안하였고, 센서융합 측면에서 성택영 등(2016)과 김봉현(2016)은 객체의 위치추적을 위해 영상센서와 GPS, LiDAR 등을 융합하여 대상 객체의 위치를 추정하는 방안을 제안하였다. 관련 기존 연구들의 경우 모니터링 관점에서 CCTV의 활용을 위한 다양한 접근 방법을 제시하고 있으나 정보 제공의 목적이 준 실시간 영상 분석에만 초점이 맞추어져 있어 모니터링 이외에 활용성이 떨어지거나 다중 센서가 요구되는 등의 단점을 가지고 있다.

본 연구에서는 기 구축된 CCTV의 활용도를 높일 수 있는 방안으로 CCTV영상에서 이동체(차량, 사람 등)들을 식별·추적하고, 이들의 위치와 관련된 정보를 지도상 공간정보로 제공함으로써 다양한 환경에서 활용할 수 있는 방안을 제안한다. 이를 위해 영상의 이동체 위치정보를 지도상의 좌표로 변환하고, 웹기반의 시각화 모듈을 통해 정밀, 항공, 일반 지도 위에 해당 객체를 공간정보로서 시각화한다.

2. 동적 객체의 위치 추적 및 시각화 방안

최근 인공지능 기술의 발달로 인해 다양한 디바이스에서 영상을 기반으로 목표 객체를 식별·추적할 수

있는 기술이 개발되고 있다. 이러한 이유로 인공지능 기술을 접목한 영상 센서인 CCTV는 다양한 분야에서 다양한 목적으로 활용될 수 있다. 이는 육안으로 확인하는 단순한 모니터링 역할을 넘어 도로 인프라 측면에서 도로 상황을 실시간으로 분석하여 관련 정보를 제공할 수 있는 중요한 센서로서의 역할을 할 수 있을 것이다. 본 연구에서 제안한 동적 객체의 위치 추적 및 시각화 방안은 데이터 로드 및 객체 추적, 좌표계 변환, Kafka¹⁾ 서버, 웹기반 위치 시각화로 구성된다. 또한, CCTV 활용 목적인 모니터링 시스템 연계를 위한 확장성을 위해 Python²⁾ 기반으로 관련 기능들을 구현하여 모듈화 하였으며, Figure 1은 제안된 방안의 개략적인 구성도를 나타낸다.

입력으로는 CCTV영상을 활용하며, 인공지능을 활용하여 영상 프레임에 포함된 동적 객체(차량, 사람, 자전거 등)를 식별하고, 객체의 최소사각형(minimum bounding box) 작성 방식으로 위치를 추적한다. 식별 및 추적된 객체에 대해서는 카메라 좌표계와 지도 간의 좌표계 변환 관계가 정립된 변환 행렬을 통해 영상좌표계로 정의된 위치를 지도좌표계로 변환한다. 또한, 다중 사용자 지원을 위해 구성된 Kafka 서버에 객체 관련 정보를 실시간으로 업로드하는 과정을 수행하며, 정보 전달의 표준화를 위해 json³⁾ 포맷을 활용한다. 이렇게 전달된 객체 정보는 다양한 디바이스 지원을 위해 제작된 웹기반의 시각화 모듈을 통해 정밀, 항공, 일반 지도 위에 시각화된다.

2.1. 동적 객체 추적

영상에서 정확한 목표 객체를 식별·추적하기 위해서는 일반적으로 객체의 형태, 색상 등의 정보들을 활용할 수 있다. 따라서 인공지능 기술이 고차원의 기능을 지원하기 전까지는 이미지 하나에서 하나의 객체를 인지하는 과정은 많은 노력과 시간이 요구되었다. 그러나, 인공지능 기술의 등장으로 하나의 이미지에

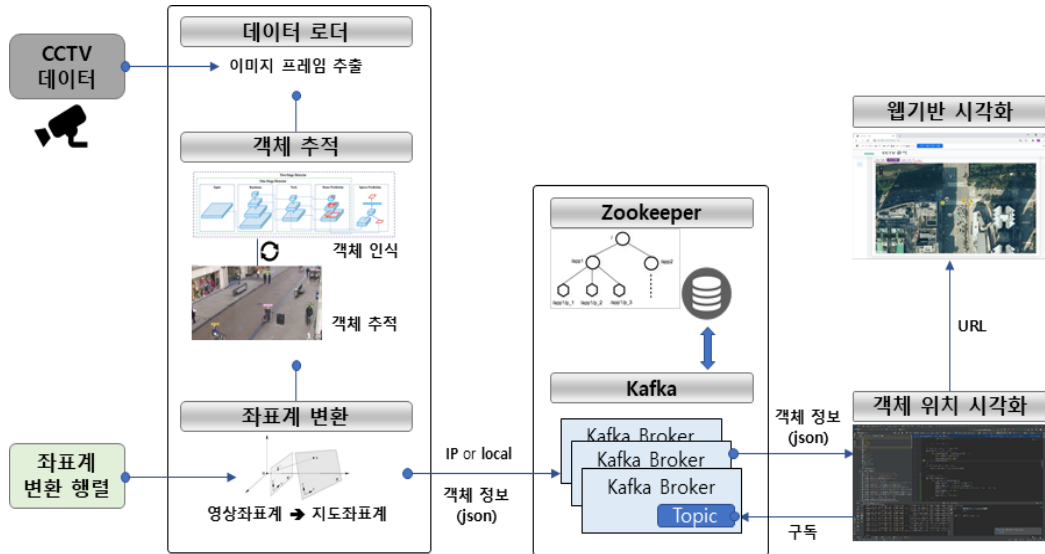


Figure 1. The identification and tracking of Dynamic Objects Using CCTV images.

서 여러 객체를 실시간에 가까운 속도로 인지하는 기술이 개발되고, 이러한 기술들은 다양한 분야에 활용되면 점차 고도화되고 있는 단계에 있다.

본 연구에서는 인공지능 기술을 활용하여 인지하고자 하는 객체를 학습한 후, CCTV영상의 프레임에서 해당 객체들을 추적하는 모델을 적용하였다. 인공지능 기술에서도 영상정보를 분석하는 기술은 제품에 탑재되어 상용화될 만큼 가장 고도화되어 있어 인지능력이 뛰어나다고 할 수 있다. 알려진 영상분석을 위한 인공지능 기술로는 Faster RCNN(Ren 2015), YOLO (Redmon 2016), SSD(Liu 2016) 등이 있다. Faster RCNN은 초기 RCNN에서 Fast RCNN을 거쳐 실시간 처리속도를 목표로 개선된 모델이다. 이는 Fast RCNN과 구조는 유사하지만, 매우 작은 region proposal network라는 회귀네트워크를 selective search에 활용하여 RCNN대비 250배, Fast RCNN대비 25배의 처리 속도 성능을 보여준다. YOLO는 이미지의 pixel 단위가 아닌 n개의 상자에 해당하는 grid 단위로 나누어 객체를 추적하는 방식이며, 실시간에 가까운 처리 성능으로 realtime detection system에 적합

한 모델이다. SSD는 다양한 크기의 feature map을 바탕으로 한 장의 이미지에서 다양한 목표 객체를 추적하는 방식을 취하는 방식으로 균형있는 정확도와 처리 속도를 보여주는 모델이라고 할 수 있다.

본 연구에서는 30fps 이상의 촬영속도를 가지는 CCTV영상 정보의 처리와 함께 객체를 추적 및 시각화하는 과정까지 실시간으로 처리하는 것으로 목표로 한다. 따라서, 객체 식별 시간이 빠르다고 알려진 YOLO 알고리즘을 채택했다. 또한, 최근 발표된 YOLOv4는 성능면에서 많은 향상을 보여주고 있다 (Bochkovskiy et al. 2020). 이러한 모델을 활용하기 위해서는 식별·추적하고자 하는 대상 객체에 대한 학습이 선행되어야 한다. 학습을 위한 대상 객체는 도로 내외의 이동체에 해당하는 차량, 사람, 이륜차(자전거 포함), 유모차 4종과 기타로 정의된 객체를 포함하여 총 5종을 정의하였고, 이에 대한 학습 데이터 셋은 해외의 다양한 학습 셋을 보유하고 있는 COCO data set⁴⁾과 국내 자율주행 데이터를 보유·제공하고 있는 LX(한국국토정보공사)의 KODAS data set⁵⁾을 활용하여 구성하였다.

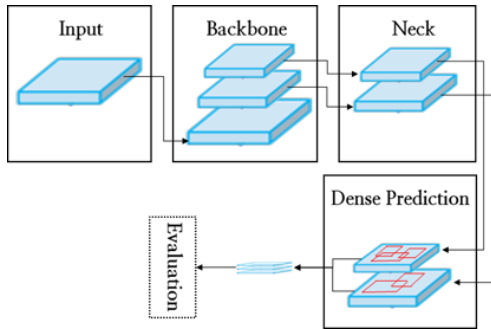


Figure 2. Architecture for YOLOv4.

학습을 위한 YOLOv4 모델은 Figure 2와 같이 Bochkovskiy et al.(2020)의 YOLOv4 아키텍처를 사용하였다. YOLOv4 모델 학습을 위한 하이퍼파라미터로는 training step을 15,000회로 하였고, batch size는 64, learning rate를 0.1로 하였다. 또한, momentum과 weight 변화량은 각각 0.9와 0.0005로 정의하여 활용하였다. 본 연구에서 설계한 모델의 성능을 확인하기 위해 공공데이터포털(data.go.kr)의 경기도 자율주행 센터 CCTV영상 이벤트 데이터⁶⁾를 사용하여 시험을 수행하였다. 개방 중인 CCTV영상 데이터를 활용하였으며, 해당 영상의 객체는 도로상황 모니터링을 위한 CCTV 위치의 특성상 대부분 차량과 사람 클래스로 주로 구분되어 있다.

Figure 3은 추적결과를 나타내며, Table 2는 영상에서 30픽셀 이상의 객체를 대상으로 IoU (intersection over Union)에 따른 AP(Average Precision) 시험 결과를 나타낸다. AP₅₀은 IoU가 0.5 이상인 객체를 정검출(True Positive)로 하였을 때의 AP를 나타내며, AP₇₅는 IoU가 0.75일 때의 AP를 나타낸다. AP_{0.5:0.05:0.95}는 IoU를 0.5부터 0.05씩 증가시켜가며 0.95까지의 AP결과들의 평균을 나타낸다. 본 시험에서는 클래스별 출연 빈도를 토대로 차량과 사람 객체를 고려하였으나, 추후에는 다양한 클래스를 갖는 여러 지역의 CCTV 영상을 활용하여 성능평가를 추가로 진행할 예정이다.

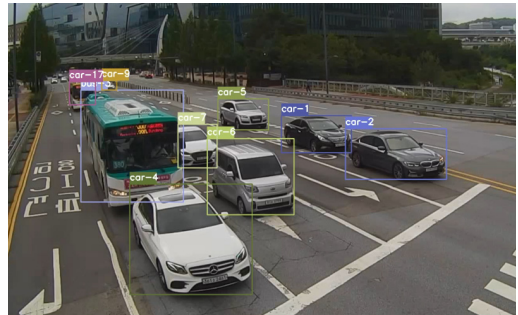


Figure 3. Result of detection using our model.

Table 2. AP results of YOLOv4

	AP _{0.5:0.05:0.95}	AP ₅₀	AP ₇₅
클래스			
car	0.54	0.66	0.56
person	0.72	0.93	0.78
평균(mAP)	0.63	0.79	0.67

목표 객체가 식별되는 이를 끊임없이 추적하는 과정도 함께 진행되어야 한다. 따라서, 과거 정보를 현재 시점에 적용하여 실제 값을 예측하는 Deep Sort (Wojke 2017) 알고리즘을 YOLOv4 뒤쪽에 붙여 연속되는 영상에서 특정 객체를 놓치지 않고 추적할 수 있도록 하였다. 이는 이전 프레임에서 식별된 객체의 특징 정보를 저장하여 이후 프레임에 적용하여 동일한 객체를 매칭하는 방식이다.

먼저, YOLOv4 기반으로 프레임 내 객체의 종류와 위치를 인지하고 그 결과를 Deep Sort에 적용하여 연속적인 프레임에서 특정 객체를 추적하는 루프가 반복적으로 수행된다. Figure 4는 연속되는 프레임에서 목표 객체의 식별 및 추적 결과를 나타낸다.

객체 위치 추적 방식으로는 최소사각형 형태를 채용하였으며, 추적된 각 객체에 대한 사각형은 향후, 시각화 정합을 위한 객체의 위치정보로 활용된다. 본 연구에서는 추적된 객체가 사람인 경우, Sankaranarayanan(2008)의 방안을 토대로 사각형 하단의 중심

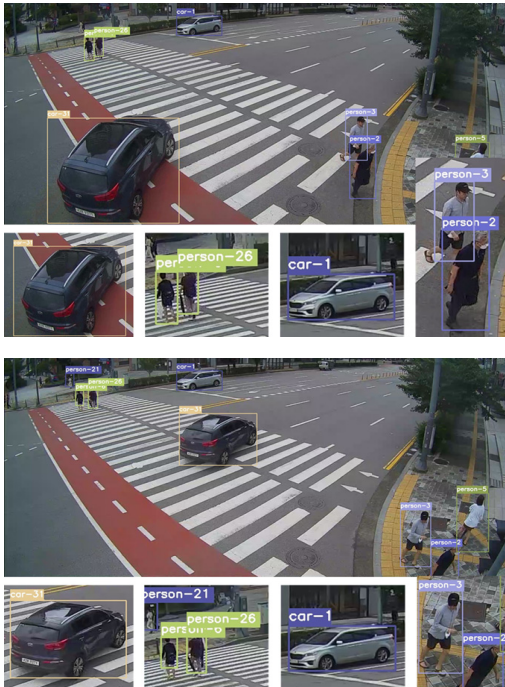


Figure 4. Results of object tracking in the CCTV images

점을 대표 위치로 활용하였으며, 차량 등의 경우 객체의 형태를 간략화한 사각형의 높이, 너비를 고려하여 하단으로부터 1/4 지점의 중심점을 대표 위치로 채택하여 활용하였다.

2.2. 좌표계 변환

본 연구에서 입력정보로 활용되는 CCTV 영상은 가로(X), 세로(Y)이 각각 1920과 1080 해상도를 가지고 있는 반면, 정밀, 항공, 일반 지도가 활용된 시각화에서는 위도(latitude) - 19~90, 경도(longitude) - 180~180의 범위를 가지는 WGS84 좌표계를 사용한다. 따라서, 입력영상에서 추적되는 객체의 정위치 매핑을 위해서는 좌표계 간의 차이(translate, rotation, scale)를 설명할 수 있는 변환 행렬이 요구된다.

CCTV영상과 시각화의 차원이 다른 두 좌표계 간의

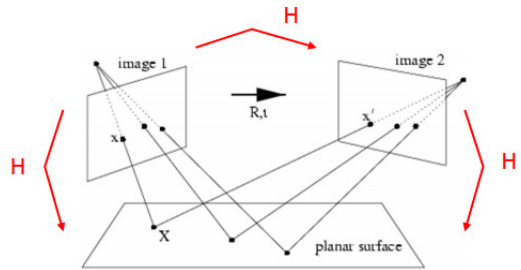


Figure 5. Definition of a homography matrix
Source: Opencv tutorial

변환 관계를 확인할 수 있는 방법으로는 CCTV의 내(intrinsic)·외부(extrinsic) 파라미터와 함께 CCTV 영상의 특정 점과 매칭되는 WGS84 좌표계 상의 점을 통해 변환 행렬을 계산하는 방법과 이미지에서 불변하는 다수의 특징점을 추출하고 WGS84 좌표계에서 이에 대응되는 특징점을 매핑하여 두 좌표계 간의 변환 관계가 정립된 변환행렬 H(Homography matrix)를 계산하는 것이다(Mehboob 2017). 전자의 경우, 위치 매칭 정확도가 높을 수 있으나 CCTV의 실 측정을 통한 내·외부 파라미터 획득이 쉽지 않다는 단점이 있다. 후자의 경우는 전자에 비해 정확도는 떨어질 수 있으나, 변환 관계 정립이 비교적 간단하고 대응 특징점의 정밀도에 따라서 위치정확도를 향상시킬 수 있다는 장점이 있다.

본 연구에서는 유용성 검증 측면에서 H 계산 방법을 채택하여 활용하였다. H는 2D-2D, 2D-3D 간 좌표계 변환에 주로 활용되는 방법으로 이는 AR(Augmented reality) 기술에서 마커와 카메라 간의 관계 정립을 위해 필수적일 활용되는 방법이다(Prince 2002). Figure 5과 같이 서로 다른 두 방향의 카메라에서 같은 점 x 를 바라볼 때 각 카메라의 영상에 맺히는 점 x_1 와 x_2 간의 이동, 회전 등이 정의된 3×3 행렬을 의미한다.

이를 본 연구의 좌표계 변환 문제에서 Figure 5의 image1을 CCTV 카메라로 정의하고, image2를 WGS84 좌표계로 정의하여 두 좌표계 간의 변환 관계를 정의

할 수 있다. 즉, CCTV 영상의 불변하는 특징점에 대한 WGS84 좌표계 대응점을 활용하여 변환 행렬을 계산하는 것이다.

이를 위해 Figure 6과 같이 객체 위치정보 시각화 대상지의 CCTV영상에서 불변하는 특징점(차선, 노면 표시, 횡단보도 등)을 정의하고, 도로의 정보를 실측하여 WGS84좌표에 제작한 정밀도로지도에서도 영상 특징점에 대응되는 점들을 정의하였다. 영상의 특징점과 정밀도로지도에서의 대응점은 일관성을 유지하기 위해 매뉴얼 방식으로 수행하였다. 여기서 영상의 일부 영역에 대한 특징점을 활용하는 경우, 특징점이 없는 영역이 왜곡되어 객체의 위치가 부정확하게 맵핑되는 경우가 발생하기 때문에 영상에서 정의하는 특징점들은 되도록 넓은 범위를 포함하도록 정의되어야 한다. 본 연구에서는 좌표계 간 변환 관계 계산을 위해 각 영상별 25개의 불변 특징점들을 정의한 후, 변환식 $P'_{wgs84} = H \times P^{T_{img}}$ 을 통해 객체의 위치를 변환하여 시각화하였다.

2.1과 2.2절을 통해 목표 객체를 식별 및 추적한 결과로는 모니터링을 위한 정보로 객체 타입, 위치, ID 등이 출력된다. 이는 하나의 객체 대한 속성들로서 용이한 전달 및 활용을 위해서는 하나의 데이터 셋으로 묶는 과정이 요구되며, 이를 위해 적용할 수 있는 포맷들은 다양하다. 본 연구에서는 많은 분야에서 활용되고 인코딩, 디코딩이 빠르고 안정적인 json 포맷을 채택하여 객체의 식별 및 추적 결과로 출력된 속성정보를 정의하였다. 이렇게 정의된 객체별 데이터 셋은 이후 Kafka 서버에 전달되고, 시각화에서는 서버로부터 관련 정보를 받아와 디코딩을 수행한 후, 배경지도에 객체 정보를 매핑하게 된다.

2.3. Kafka 서버

자율주행을 위한 통합센터 또는 도로교통 관제센터 등에서는 도로상황 모니터링을 위해 관리 구역에 여

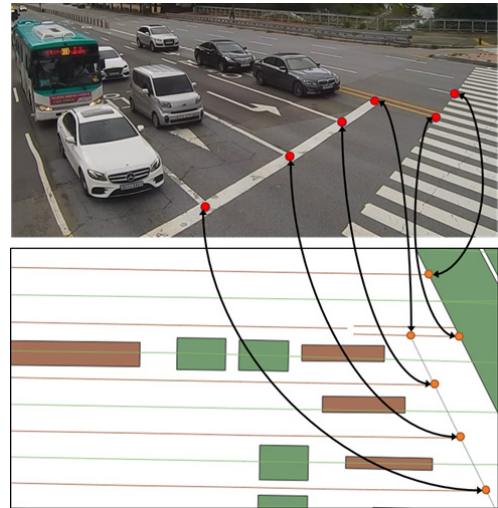


Figure 6. Definition of corresponding points for two coordinate systems

러 대의 CCTV를 설치·운영하고 있다. 30fps의 촬영 속도를 가지는 여러 대의 CCTV 정보를 실시간으로 분석하고, 이를 여러 사용자에게 전달하는 과정은 하드웨어 및 소프트웨어 측면에서 많은 리소스가 요구되며, 세부적으로는 빠른 데이터 적재와 제공이 필요 조건이라 할 수 있다. 이를 위해 본 연구에서는 분산처리 환경에 특화된 Kafka를 채택하여 이를 객체 정보 적재와 시각화 정보 전달에 활용하였다. Kafka는 실시간으로 데이터를 저장 및 처리할 수 있는 분산 데이터 스트리밍 플랫폼으로 대규모 데이터를 1:1 데이터 교환뿐만 아니라 1:N, N:1 간의 동시 전달이 가능하다. 즉, 여러 소스에서 데이터 스트림을 처리하고 여러 사용자에게 전달하는 오픈소스 기반의 솔루션이라 할 수 있다.

Kafka는 Figure 7과 같이 크게 Producer, Consumer, Kafka cluster, Topic으로 구성되는데, Producer는 Topic 노드에 데이터 쓰기 요청을 보내 지정된 토픽에 메시지를 작성(발행)하는 주체이며, Consumer는 생성된 Topic에 접근(구독)하여 데이터를 받는 클라이언트에 해당한다. Kafka cluster는 생성된 Topic

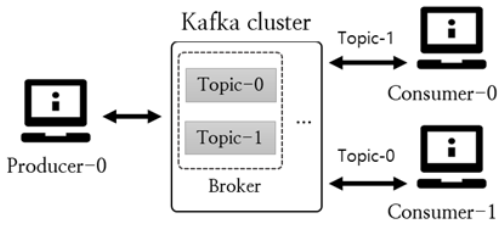


Figure 7. Definition of the kafka structure

과 관련 데이터를 파티션 단위로 분산 관리하는 역할을 수행한다. Kafka가 실시간 데이터 처리에 적합한 이유는 이들 구성 요소들이 분리되어 작동되기 때문에 장애 발생 의존성이 낮고, 메시지 큐가 아닌 Topic을 통해 직접 데이터를 전송하는 방식으로 이루어지기 때문이다.

이를 토대로 본 연구에서는 데이터 전달·관리를 위한 Kafka cluster를 구성하고 공유를 위한 Topic 노드를 정의하여 관련 파티션을 구성한 후, Consumer에 해당하는 시각화 모듈이 해당 Topic 노드를 구독할 수 있도록 설정하였다. 따라서, Producer에 해당하는 목표 객체의 식별·추적 모듈에서 Kafka에 정의된 Topic으로 결과를 전달하면 Kafka cluster에서 해당 Topic 노드에 관련된 파티션에 데이터가 적재되고, 이와 동시에 구독상태에 있는 시각화 모듈에서 해당 데이터를 받아갈 수 있게 되는 것이다. 이러한 Kafka기반의 데이터 전달 방식은 요청에 의한 전달 방식이 아니기 때문에 데이터 전달 지연을 최소화 할 수 있고, 관련 Topic 노드를 구독하는 것만으로 업데이트 즉시 데이터를 최신화할 수 있어 본 연구에서 목표로 하는 실시간 다중 사용자 지원을 달성할 수 있는 기반이 된다.

2.4. 웹기반 시각화

최근 스마트폰 등 단말기의 사양 향상과 정보 전달의 방식 변화로 인해 정보 전달을 위한 시각화 방법도 기존의 단일 소프트웨어 방식에서 다양한 플랫폼에서

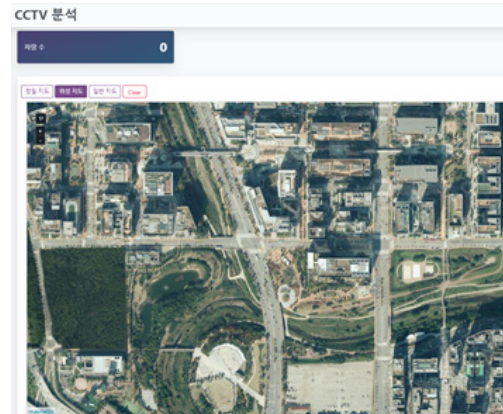


Figure 8. Web-based object visualization

활용할 수 있는 웹 방식으로 변화하는 실정이다. 본 연구에서는 웹개발이 용이한 Spring framework를 기반으로 동적인 웹기반 시각화 기능을 구현하였다. 이는 일반 웹개발 언어에 비해 다양한 서비스들을 지원하며, 특히 데이터베이스 처리 라이브러리와 호환성이 뛰어나 기존 라이브러리들을 통합·운용할 수 있는 환경을 제공해 준다. 본 연구의 CCTV영상내 객체의 시각화 환경은 Figure 8와 같이 구성되어있다.

시각화를 위한 웹 모듈은 간단한 구조로 이루어져 있는데, 베이스 지도로 정밀지도와 항공지도, 그리고 일반지도를 지원하고 있으며 사용자는 베이스맵 선택 등의 기능을 활용할 수 있다. 해당 웹 모듈은 2.3절의 Cosumer에 해당하는 Client로 Kafka 서버가 구동중인 상태에서 관련 Topic에 접근하여 객체들의 위치 및 관련 정보를 모니터링할 수 있는 기능을 제공한다. Figure 9는 1 CCTV, 3 CCTV 대해 제안된 방안을 적용한 결과를 각각 나타낸다. 활용성 측면에서 주간 촬영된 1 CCTV영상을 토대로 대상 객체를 식별·추적한 결과 CPU 3.6Ghz, 32Gb RAM, GPU 1080ti 환경에서 객체 시각화까지의 성능은 18(±5)fps의 처리 속도를 보였다.

이는 CCTV 촬영속도인 30fps에 못 미치는 수준이나, 현 GPU 사양을 감안했을 때 고사양의 GPU를 활용

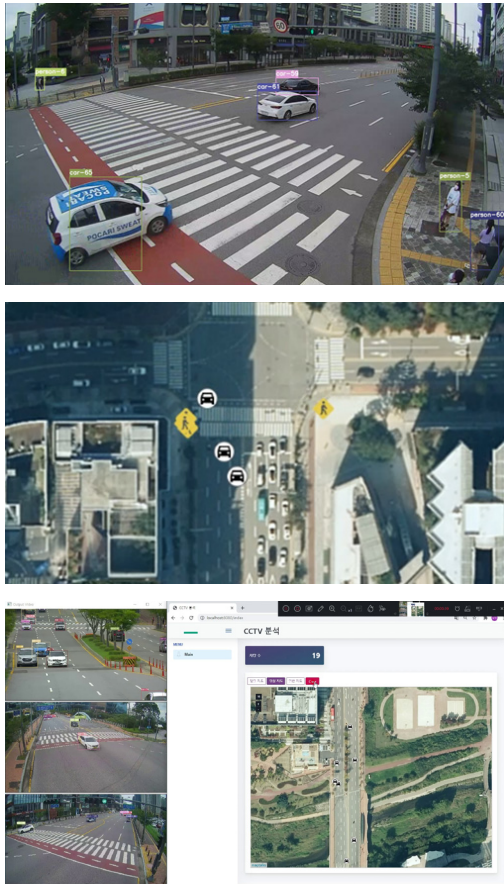


Figure 9. Results of the proposed method

할 경우 30fps 이상의 처리 속도를 보일 수 있을 것이다. 영상분석 기술 특성상 임계치 내에서 영상의 크기 조절은 결과에 많은 영향을 미치지 않기 때문에 영상 크기를 조절하여 처리속도를 올리는 것도 다른 방안이 될 수 있다.

또한, 실시간 다중 사용자 지원 점검을 위해 CCTV와 시각화 모듈을 각각 8대까지 교차 증가시켜 수행한 시험에서는 객체 추적에서 Kafka 서버 적재, 시각화까지의 처리 속도가 각 1대씩을 활용한 성능인 18(±5) fps를 유지함을 확인하였고, 이를 토대로 영상분석 파트에서 데이터 처리를 위한 대부분의 리소스가 활용됨을 알 수 있었다. 보유 데이터량의 한계로 수십 대에

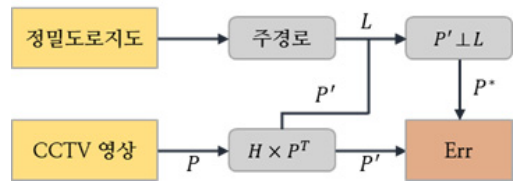


Figure 10. Process for error estimation

달하는 CCTV 영상까지는 적용해 보지 못했으나, 본 연구에서 구축한 Kafka 시스템의 경우 제한된 데이터 전달 포맷 및 객체 속성을 따른다면 실시간 다중 지원을 위한 서버 시스템으로 활용이 가능할 것으로 사료된다.

3. 위치 정합도 평가

제안된 방안에서는 CCTV영상에서 차량, 사람 등에 대한 이동체를 공간좌표계 상의 정위치에 매핑하는 부분이 매우 중요한 요소이다. 따라서, 위치 정합도에 대한 검증이 요구되며, 이를 위해 본 연구에서는 CCTV영상에서 주경로의 임의의 점과 이를 변환한 WGS84 변환점의 일치 정도를 검토하기 위해 Figure 10과 같은 과정을 토대로 좌표계 변환 정확도를 평가하였다.

먼저, 실측을 통해 구축된 정밀도로지도 구조에서 Figure 11에서와 같이 차선과 차선 사이에 도로의 합류, 분기, 교차로 등에서 차량이 어느 방향으로 운행해야 되는지 안내하는 주경로(L)를 얻은 다음, CCTV영상에서 차선을 찾고, 차선과 차선 사이의 중앙에 위치하는 영상기반의 주경로를 계산하여 임의의 n 개 점 P 을 샘플링한다. 샘플링은 정확도 확보를 위해 영상처리 라이브러리인 OpenCV를 활용하여 차선 추출과 주경로 계산 등을 진행하였다(Farag 2018).

이렇게 CCTV영상에서 샘플링된 점들에 대해 2.2절의 좌표계 변환 방안을 활용해 정밀도로지도의 공간 좌표계의 점(P')으로 변환하고, P' 에서 L 과 가장 거

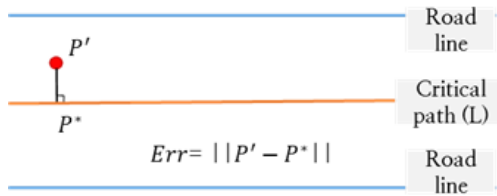


Figure 11. Definition of error estimation

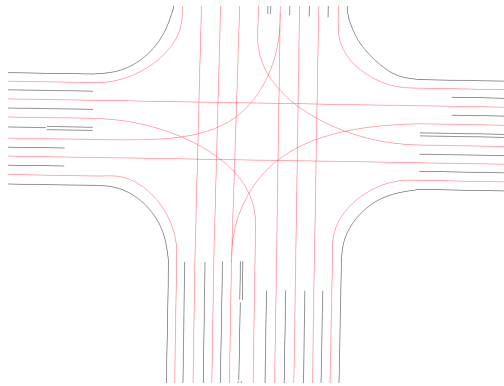


Figure 12. Road line and critical path in the HD map

리가 가깝고 수직인 위치에 있는 수직 교차점(P^*)를 찾는다. 좌표계 변환이 오차 없이 잘 이루어졌다면 계산된 점 P^* 는 정밀도로지도 주경로 L 위에 위치하게 된다.

즉, Figure 11의 Err 가 '0'에 가까울수록 위치 정확도가 높다고 할 수 있다. Figure 12는 차선(검정색 선)과 L (빨간색 선)로 정의된 정밀도로지도의 구조를 나타내며, Figure 13은 정밀도로지도 L 과 대응되는 CCTV영상의 주경로에서 평가를 위해 샘플링한 임의의 점 P 를 나타낸다. 본 연구에서는 26개의 임의의 점을 샘플링 후(Figure 13 참조), Figure 10의 과정을 통해 제안된 방안의 위치 정확도를 평가하였다.

Table 3은 변환된 P' 와 수직 교차점 P^* , 그리고 두 점 간의 거리를 비교한 결과를 나타낸다. 거리 비교 결과, 작게는 0.02m에서 크게는 0.48m 까지 오차가 발생하였고, Avg (Average)와 RMS(Root Mean Square)

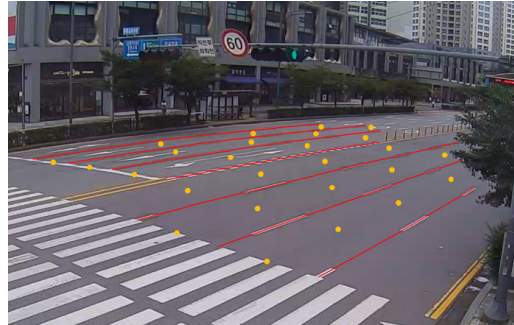


Figure 13. Sample points of a CCTV image

Table 3. Results of the evaluation

n	P'	P^*	Err (m)
1	37,39810236995146	37,39810236995146	0.27
	127,11284371695545	127,11284458460997	
2	37,398056022983745	37,398056022983745	0.06
	127,11284453412071	11284459860855	
3	37,398014448299726	37,39801444225397	0.41
	127,11284392441756	127,1128446110664	
4	37,3981003578181	37,39810034847469	0.48
	127,11287948619685	127,11288081438913	
5	37,398053366157214	37,39805336615721	0.31
	127,11288048901872	127,11288084370179	
⋮			
22	37,39805187414329	37,398051865349494	0.36
	127,11291959905732	127,11291840016989	
23	37,39799629735647	37,39799629149394	0.37
	127,11291879882195	127,11291837671978	
24	37,39804721105009	37,39804722277515	0.24
	127,11296406345524	127,11296141945408	
25	37,39790503296632	37,397905021241264	0.44
	127,11296219917062	127,11296124944073	
26	37,39794793496272	37,397947934962716	0.02
	127,11291813053069	127,11291835477248	
Avg			0.15
RMS			0.19

는 각각 0.15m와 0.19m의 오차를 보였다. CCTV와 가까운 영역에서 샘플링된 점의 경우 비교적 낮은 오차를 보였는데, 이는 두 좌표계에 대한 변환 행렬 계산 시 카메라 시야에서 가깝고 왜곡이 적은 영역에서 얻은 특징점은 잘 반영되는 반면, 카메라 시야에서 멀리

있는 영역은 심한 왜곡으로 특징점이 정확히 반영되지 않아 생기는 현상으로 판단된다. 즉, 변환 행렬 계산을 위한 특징점은 왜곡이 심한영역에서 보다 세밀하고 많은 정보를 얻을 수 있는 방향(CCTV에서 거리가 멀어질수록 특징점 개수 증가 등)으로 정의할 필요가 있으며, 이를 통해 더 높은 위치 정합도를 기대할 수 있을 것이다.

4. 결론

본 연구에서는 도로 인프라에서 도입 및 활용되는 센서 중에서 많은 구간에 구축은 되었으나 단순한 모니터링 이외에 활용도가 적은 CCTV를 보다 적극적으로 활용할 수 있는 방안으로써 영상에서 인공지능 기술을 통해 목표 객체를 추적하고, 이를 시각화하는 방안을 제안하였다. 제안된 방안에서는 추적을 위한 차량, 사람 등의 목표 객체를 정의하고, 이를 식별·추적할 수 있는 인공지능 모델과 다중 사용자 지원을 위한 실시간 데이터 제공 서버 및 정밀, 항공, 일반 도로를 베이스 맵으로 하는 시각화 기능을 구현하였다. 또한, 제안된 방안의 유효성 검증을 위해 위치 정합도 평가 절차를 정립하여 평가를 수행하고, 이를 통해 제안된 방안의 적용 가능성과 유효성을 확인하였다. 제안된 방안을 통해 CCTV를 다양한 분야에서 활용할 수 있는 기반을 마련하였고, 이를 위한 프로세스는 많은 영상기기를 활용한 제품 및 연구에서 다양하게 될 수 있을 것으로 사료된다.

그러나, 좌표계 변환 관계 정립에 많은 영향을 받는 만큼 보다 정확한 정립 방안이 요구되며, CCTV 활용 고도화를 위한 기초연구로써 객체들 간의 가림, 환경(날씨)등과 같은 조건은 고려하지 않은 한계를 가지고 있다.

향후 연구에서는 시각화 처리 속도와 위치 정합도 향상을 위해 지지체와의 협의를 통해 더 많은 다양한 형태의 CCTV 데이터를 확보하고, 이를 토대로 객체 식

별·추적 모델과 좌표계 변환 방안을 개선하고자 한다.

- 주1. LinkDin에서 개발된 분산 메시징 시스템으로 2011년에 오픈소스로 공개되었으며, 대용량의 실시간 로그 처리에 특화된 아키텍처 설계를 통하여 기존 메시징 시스템보다 우수한 TPS(Transaction Per Second) 성능을 보임.
- 주2. 1991년 귀도 반 로섬(Guido van Rossum)이 발표한 객체지향 고급 프로그래밍 언어로, 플랫폼에 독립적이며 인터프리터식, 객체지향적, 동적 타이핑기반의 대화형 언어임. 이는 개방형, 공동체 기반 개발 모델을 가지고 있으며, C언어를 기반으로 하고 있음.
- 주3. JavaScript Object Notation의 축약어로 속성-값 쌍 또는 키-값 쌍으로 이루어진 데이터 오브젝트를 전달하기 위한 텍스트 기반의 개방형 표준 포맷이며, 인터넷에서 자료를 주고받을 때 주로 사용하는 표현 방법으로 알려져 있음
- 주4. Object detection, Segmentation, keypoint detection 등을 위한 데이터 셋으로, 매년 다른 데이터 셋을 전 세계의 여러 대학/기업이 참가하는 대회 및 연구에 사용하고 있음
- 주5. 한국국토정보공사(LX)에서 구축·공개 하고 있는 자율주행에 특화된 도로 환경 데이터(Lidar, Rader, image 등)로 객체 식별 추적을 위한 학습데이터를 구축하여 제공하고 있음

감사의 글

본 연구는 산업통상자원부와 한국산업기술진흥원의 “새만금지역 상용차 자율주행 테스트베드 구축사업”(과제번호 P0013841)으로 수행된 연구결과입니다.

참고문헌

References

- 김봉현. 2016. 위치 측위 기술을 이용한 영상 추적 시스템 설계. 한국디지털정책학회. p.143-148.
- Kim BH. Design of Image Tracking System Using Location Determination Technology. The Society of Digital Policy&Management. p.143-148.
- 김진태. 2018. 자율주행차의 도로주행을 위한 운행체계 및 교통인프라 연구개발 기획연구. 도로교통

- 공단. 2018-12.
- Kim JT. 2018. A Study on the R&D fo the Operation System and Transportation Infra-structure for Road Driving of Self-driving Cars. The Road Traffic Authority. 2018-12.
- 박종록, 김한해. 2019. 자율주행기술. 한국과학기술기 획평가원. 2019(16).
- Park JL, Kim HH. 2019. Autonomous driving Technology. 2019(16).
- 서홍덕, 김의명. 2020. 스테리오 CCTV영상에서 딥러 닝을 이용한 교통량 추정. 한국측량학회. 38(3): 269-279.
- Seo HD, Kim EM. 2020. Estimation of Traffic Volume Using Deep Learning in Stereo CCTV Image. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*. 38(3):269-279.
- 성택영, 권기창 et al. 2016. 단일 카메라와 GPS를 이 용한 영상내 객체 위치 좌표 추정 기법. 한국멀티 미디어학회. 10(2):112-121.
- Seung TY, Kwon GC et al. 2016. An Estimation Method for Location Coordinate of Object in Image Using Single Camera and GPS. *Journal of Korea Multimedia Society*. 10(2):112-121.
- 양인철, 전우훈 et al. 2019. 교통 데이터 수집을 위한 객체 인식 통합 프레임워크 개발. 한국ITS학회. 18(6):191-201.
- Yang IC, Jeon WH. 2019. Development of an Integrated Traffic Object Detection Frame-work for Traffic Data Collection. *The Korea Institute Of Intelligent Transport Systems*. 18(6):191-201.
- ACM. 2017. American Center For Mobility: <https://www.acmwillowrun.org/>.
- Kiela K, et al. 2020. Review of V2X-IoT Standards and Frameworks for ITS Applications. *MDPI applied sciences*. 10(12):2-23.
- Bezzina D, Sayer J. 2015. Safety Pilot Model Deployment:Test Conductor Team Report. NHTSA; [accessed 2015 Jun]. <http://www.nhtsa.gov/>.
- Bochkovskiy A, et al. 2020. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.
- CV Pilot. 2017. Connected Vehicle Pilot Deploy-ment Program: United States Department of Transportation(ITS): <https://www.its.dot.gov/pilots/>.
- Farag W, Saleh Z. 2018. Road lane-lines detection in real-time for advanced driving assistance systems. IEEE on 3ICT. p.1-8.
- Kotsi A. 2020. Overview of C-ITS Deployment Projects in Europe and USA: [accessed 2020 Oct 14]: <https://arxiv.org/abs/2010.07299/>.
- Liu W, et al. 2016. Ssd: Single shot multibox detector. European conference on computer vision. p.21-37.
- Mehboob Fozia, et al. 2017. Glyph-based video visualization on Google Map for surveillance in smart cities. Journal on Images and Video Processing. 2017-28.
- Prince S.J.D, et al. 2002. Augmented reality camera tracking with homographies. *IEEE Computer Graphics and Application*. 22(6):39-45.
- Ren S, et al. 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Projpsal Networks. Neual Information Processing Systems (NIPS).
- Redmon J, et al. 2016. You only look once: Unified, real-time object detection. Proceedings of the

IEEE conference on CVPR, p.779-788.	on ICIP, p.3645-3649.
Sankaranarayanan A, et al. 2008. Object Detection, Tracking and Recognition for Multiple Samrt Cameras. <i>Proceedings for the IEEE</i> , 96(10): 1606-1624.	2021년 4월 23일 원고접수(Received)
Wojke N, et al. 2017. Simple online and realtime tracking with a deep association metric. <i>IEEE</i>	2021년 6월 7일 1차심사(1st Reviewed)
	2021년 6월 21일 2차심사(2nd Reviewed)
	2021년 6월 28일 게재확정(Accepted)

초 록

국내·외적으로 수행되고 있는 다양한 C-ITS 관련 도로 인프라 구축 사업들은 다양한 센서 기술들을 융합적으로 활용하고 있으며, 도로 인프라의 효율성과 신뢰성을 높이기 위해 센서 관련 기술 향상에 많은 노력을 하고 있다. 최근에는 인공지능 기술의 발전으로 영상정보를 수집하는 CCTV의 역할은 더욱 중요해지고 있다. CCTV는 현재 도로 상태 및 상황, 보안 등의 이유로 많은 양이 구축되어 운영되고 있으나, 단순한 영상 모니터링에 주로 활용되고 있어 자율주행 측면에서 센서들에 비해 활용도가 부족한 실정이다. 본 연구에서는 기구축된 CCTV영상에서 이동체(차량·사람 등)들을 식별·추적하고, 이들의 정보를 다양한 환경에서 활용할 수 있도록 분석·제공하는 방안을 제안한다. 이를 위해 Yolov4와 Deep sort 알고리즘을 활용한 이동체 식별·추적과 Kafka 기반의 실시간 다중 사용자 지원 서버 구축, 영상과 공간 좌표계 간의 변환 행렬 정의, 그리고 정밀도로지도, 항공맵 등을 활용한 맵기반 이동체 시각화를 진행하였으며, 유용성을 확인하기 위한 위치 정확도 평가를 수행하였다. 제안된 방안을 통해 CCTV가 단순한 모니터링 역할을 넘어 도로 인프라 측면에서 도로 상황을 실시간으로 분석하여 관련 정보를 제공할 수 있는 중요한 센서로서의 역할을 할 수 있음을 확인하였다.

주요어 : 자율주행, 도로 인프라, CCTV, 인공지능, 좌표계 변환