



딥러닝 기반 의미론적 동적 격자지도를 활용한 라벨 기반 객체 추적 성능 향상 기법

원종진¹⁾ · 안경재¹⁾ · 강연식^{*2)}

국민대학교 자동차모빌리티대학원¹⁾ · 국민대학교 자동차공학과²⁾

Label-Guided Multi-Object Tracking Using Deep Learning-Based Semantic Dynamic Occupancy Grid Map

Jongjin Won¹⁾ · Kyungjae Ahn¹⁾ · Yeonsik Kang^{*2)}

¹⁾Graduate School of Automobile and Mobility, Kookmin University, Seoul 02707, Republic of Korea

²⁾Department of Automotive Engineering, Kookmin University, Seoul 02707, Republic of Korea

(Received 4 July 2025 / Revised 2 October 2025 / Accepted 10 October 2025)

Abstract : This paper presented a method for improving multi-object tracking performance by incorporating object label information derived from a Semantic Dynamic Occupancy Grid Map (SDOGM). The SDOGM is continuously updated using deep learning - based perception algorithms, but its outputs may still include missed detections, false positives, measurement noise, and temporary inaccuracies in label probability and static/dynamic state classification. To address these challenges, this study proposed an Extended Kalman Filter (EKF) framework that employs label-specific motion models. In addition, the SDOGM-derived object labels and motion states were further refined, and issues such as incorrect heading estimation due to vehicle reversal were mitigated through a dual-model prediction approach. The suggested method was evaluated using the nuScenes dataset and an internal experimental dataset, confirming its improved accuracy in state estimation and label classification.

Key words : Autonomous driving(자율주행), Semantic dynamic occupancy grid map(의미론적 동적 격자지도), Multi-object tracking(다중 객체 추적), Extended kalman filter(확장 칼만 필터)

1. 서론

최근 자율주행 기술의 발전으로 다양한 지역에서 상업적 서비스를 제공하는 자율주행 차량이 점차 증가하고 있다.^{1,3)} 이에 따라 자율주행 기술의 신뢰성과 안전성 확보가 주목받고 있으며, 인지 및 추적 기술의 정밀도와 강건성이 중요한 연구 주제로 부상하고 있다. 특히 도심 환경은 다양한 객체가 복잡하게 상호작용하며 존재하여, 인지 알고리즘의 높은 강건성이 요구된다. 하지만 인지 알고리즘은 미인식, 오인식, 측정 오차 등의 문제가 동반된다.⁴⁾ 이러한 문제를 추적 알고리즘을 통해 보정함으로써 객체 정보 추정의 신뢰도를 향상시킬 수 있다. 하지만 동적 객체의 라벨에 따라 거동 특성이 다르며 이는

객체를 추적하는 과정에서 중요한 요소로 작용할 수 있다. 하지만 기존의 추적 알고리즘은 객체의 라벨 정보를 활용하지 않는다.^{5,9)} 그래서 인지된 객체의 강건성 확보를 위해, 객체의 라벨에 특화된 추적 알고리즘에 관한 추가적인 연구가 필요하다. 그래서 본 논문에서는 라벨에 따른 다중 객체 추적 기법의 성능을 개선하기 위한 새로운 접근 방식을 제안한다.

기존의 다중 객체 추적 연구는 다음처럼 구분될 수 있다. 첫째, 객체 라벨 정보를 고려하지 않고 위치, 속력 등 물리적 정보만을 기반으로 추적을 수행한 연구들이 존재한다.¹⁰⁻¹³⁾ 이러한 방식은 계산 효율성 측면에서는 유리하지만, 차량과 보행자 등 객체의 라벨 움직임 특성을

*Corresponding author, E-mail: ykang@kookmin.ac.kr

[†]This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

반영하지 못하여 복잡한 도시 환경에서는 추적 정확도 저하가 발생한다.

둘째, 객체 라벨 정보를 일부 활용한 연구들이다.¹⁴⁻¹⁵⁾ 대부분 라벨은 초기 필터 설정이나 단일 모션 모델 선택에만 사용되며, 시간에 따른 라벨 변화나 신뢰도 관리는 이루어지지 않는다. 예를 들어 Interacting Multiple Model (IMM)¹⁶⁾ 기반 구조에서는 라벨별 모션 모델을 병렬적으로 운용할 수 있지만, 많은 연산시간과 라벨 일관성 검증이나 추적 기반의 라벨 정정 과정은 포함되지 않는다.

셋째, 동적 격자지도의 인지 데이터를 활용한 다중 객체 추적 연구들은 주로 객체의 점유 상태와 속도를 추정하는 데 초점을 맞추고 있으며, 시맨틱 정보는 활용하지 않거나 단순한 정적 라벨을 보조적으로 사용하는 수준에 머무르고 있다.¹⁷⁻¹⁹⁾ 이러한 방식은 움직임 탐지나 경로 예측에는 효과적일 수 있으나, 라벨 정보가 추적 과정에서 적극적으로 반영되지 않기 때문에 라벨 일관성의 유지나 라벨에 따른 거동 추정에 한계가 존재한다.

이에 본 논문에서는 의미론적 동적 격자지도 기반의 인지 결과를 활용하고, 객체 라벨에 따라 차별화된 모션 모델과 공분산을 적용하는 확장 칼만 필터 기반 다중 객체 추적 기법을 제안한다. 또한, Tracklet manager와 Dual prediction model을 도입하여 도심 환경에서 자주 발생하는 미인식 및 오인식 상황을 개선하고 차량 방향 오인식 문제를 효과적으로 보완하고자 하였다. 더불어 객체의 동/정적 상태 전환과 라벨 정보를 추적 알고리즘 내에서 추가로 보정함으로써, 복잡한 도심 환경에서도 안정적이고 강건한 객체 추적이 가능함을 본 논문에서 제시한다.

본 논문의 구성은 다음과 같다. 제2장에서는 의미론적 동적 격자지도 기법을 설명하고, 제3장에서는 객체 라벨 정보를 기반으로 한 다중 객체 추적 알고리즘 프레임워크를 제시한다. 제4장에서는 nuScenes 데이터셋,²⁰⁾과 실

차 주행을 통해 추가로 수집한 데이터셋을 활용하여 실험을 수행하고, 단일 모델 기반 추적 기법과의 성능 평가 결과를 제시한다.

2. 의미론적 동적 격자지도

기존의 동적 격자지도는 차량 주변 환경을 격자 형태로 분할하여 각 격자 셀의 점유 확률과 속도 정보를 추정하는 방식으로, 자율주행 분야에서 널리 연구되어 왔다.²¹⁻²⁵⁾ 이를 통해 주변 객체가 정적인지 동적인지를 직관적으로 파악할 수 있으며, 환경 변화가 빈번한 복잡한 도심 환경에서도 비교적 안정적으로 객체 추적과 상태 추론을 수행할 수 있다. 특히, Bird's Eye View(BEV) 기반으로 주변을 인지함으로써 도심지와 같은 복잡한 환경을 보다 효과적으로 해석할 수 있다. 그러나 일반적인 동적 격자지도는 객체 라벨 정보를 제공할 수 없어, 차량, 보행자, 자전거 등 다양한 유형의 객체를 구분하는 데 한계가 존재한다. 이러한 한계를 극복하기 위해 객체의 의미 정보를 함께 추정하는 의미론적 동적 격자지도 기법이 제안되었다.

본 논문에서 인지 값 추출을 위해 사용된 의미론적 동적 격자지도의 격자 셀에 라벨 확률을 부여하기 위해 두 가지 딥러닝 모델을 활용하였다. 첫 번째는 BEVFusion²⁶⁾으로, 객체의 라벨, 위치, 형상을 예측하는 모델이며, 두 번째는 2DPASS²⁷⁾로, LiDAR 포인트에 라벨 값을 부여하여 위치와 라벨 정보를 예측하는 모델이다. 이러한 이유로, 의미론적 동적 격자지도는 본 논문에서 제안하는 라벨 기반 추적 기법과 높은 호환성을 갖는다.

본 장에서는 먼저 BEVFusion과 2DPASS의 네트워크 구조 및 출력 정보를 설명한 후, 이 두 모델의 결과를 Dempster-shafer²⁸⁾ 이론에 따라 통합하여 시맨틱 정보를

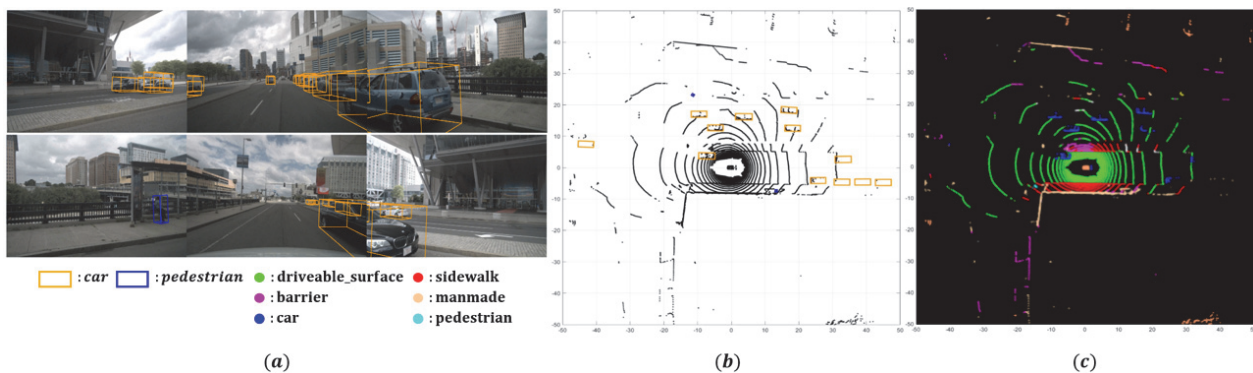


Fig. 1 Input sources for the semantic dynamic occupancy grid map include: (a) object detection results from BEVFusion applied to multi-view camera images, (b) BEV-level object representations produced by BEVFusion, and (c) semantic segmentation outputs from 2DPASS applied to LiDAR point clouds

격자 지도에 반영하는 과정을 다룬다. 또한, BEVFusion과 2DPASS에서 제공하는 다양한 라벨은 차량, 바이크, 보행자, 기타의 4개 클래스로 재분류하였다.

2.1 BEVFUSION

BEVFusion은 BEV에 기반한 다중 센서 융합 기법으로, 자율주행 차량이 주변 환경을 종합적으로 인식할 수 있도록 설계된 딥러닝 기반 인지학습 모델이다. 일반적으로 카메라 데이터는 2D 합성곱 신경망²⁹⁾을 통해 특징을 추출한 뒤, 픽셀 단위 깊이 추정을 바탕으로 BEV 공간으로 투영하는 과정을 거치며, LiDAR 데이터는 3D 스파스 컨볼루션으로 포인트 클라우드의 구조적 특징을 추출한다. 이렇게 서로 다른 센서에서 생성된 BEV 기반의 특징 맵들은 융합 모듈에서 공간적으로 정렬 및 결합되어 통합된 표현이 합성된다. 이 과정에서 센서 간 보완 정보를 고려한 다중 센서 퓨전을 수행한다. Fig. 1의 (a)는 도심 환경에서 다중 카메라 이미지 상에 탐지된 객체 정보를 시각화한 결과로, 객체 라벨별로 구분된 색상이 적용되어 있으며, 각 객체의 중심 위치, 형상, 라벨을 2D 이미지에 투영하여 표현하고 있다. Fig. 1의 (b)는 동일한 시점의 인지 결과를 BEV 관점에서 시각화한 장면으로, 자율주행 차량을 기준으로 반경 50 m 내의 객체들이 효과적으로 탐지되었음을 확인할 수 있다.

2.2 2DPASS

2DPASS는 라이다 데이터를 기반으로 한 시맨틱 분할 기법으로, 기존 라이다 시맨틱 기법이 단일 센서의 한계로 인해 발생하는 문제를 보완하기 위해 설계되었다. 학습 과정에서 RGB 이미지와 라이다 데이터를 함께 사용하여, 카메라 이미지의 풍부한 시각적 정보를 포인트 클라우드 학습에 활용한다. 이 기법은 RGB 이미지 데이터를 포인트 클라우드 정보에 투영하여 함께 학습하는 멀티 센서 융합 방식을 적용하며, 멀티 스케일 퓨전-싱글 지식 증류 전략을 활용해 이미지와 포인트 클라우드의 특징을 효과적으로 결합한다. 이를 통해 2DPASS는 더욱 정교한 시맨틱 분할을 수행할 수 있으며, 훈련이 완료된 후에는 RGB 이미지를 사용하지 않고도 높은 성능을 유지할 수 있어 실시간 응용에서도 추가적인 계산 부담을 최소화할 수 있다. Fig. 1의 (c)는 2DPASS를 통해 얻어진 시맨틱 분할 결과를 보여준다. 각 라이다 포인트는 분류된 라벨에 따라 색상이 부여되어 있으며, 차량은 파란색, 보행자는 하늘색, 도로 표면은 붉은색으로 시각화되어 있다. 특히 주변 차량, 도로, 인도, 사람 등 다양한 클래스가 명확히 분리되어 표현되어 있어, 본 기법이 복잡한 도심 환경에서도 효과적인 객체 인식 및 시맨틱 분할을 수

행함을 확인할 수 있다.

2.3 시맨틱 정보 융합

BEVFusion과 2DPASS로부터 독립적으로 산출된 라벨 정보는 서로 상이한 불확실도를 가진 채 격자 셀 단위로 수집된다. 본 논문에서는 이러한 불확실도를 Dempster-shafer 이론을 활용하여 융합함으로써, 각 셀의 라벨에 대한 신뢰도를 계산하고 이를 기반으로 의미론적 동적 격자지도를 구성한다. 아울러, 제한한 의미론적 동적 격자지도는 필터 구조 내에 라벨 확률 정보를 직접 반영함으로써, 두 인지 알고리즘 간 신뢰도를 유연하게 조절할 수 있는 특징을 가진다.³⁰⁾

의미론적 동적 격자지도의 격자 상태는 점유 확률, 비점유 확률, 동적 상태, 속도, 차량 확률, 바이크 확률, 사람 확률, 기타 확률, 불확실성으로 정의된다.

의미론적 동적 격자지도에서는 센서 데이터가 들어올 때마다 격자 셀의 상태가 지속적으로 갱신되며, 파티클 필터³¹⁾를 이용하여 점유 및 동적 상태를 예측한다. 파티클 필터를 적용하여 격자의 동적 특성을 추정하며, 각 파티클은 특정 위치, 속력, 가중치, 라벨 정보를 포함한다. 특정 격자 셀 내에서 가중치의 합이 높을수록 해당 격자의 점유 확률이 증가하며, 이는 최종적으로 Dempster-shafer 이론과 결합되어 보다 정교한 상태 추론이 가능해진다. 파티클 기반 예측 과정에서는 T 시간대의 측정된 격자에 대해 Constant Velocity(CV) 모션 모델을 적용하여 T+1 시간대의 격자 상태를 구성한다. 이때, 비점유 상태는 시간이 지남에 따라 Discount factor를 통해 신뢰도가 점진적으로 감소하도록 모델링된다. 또한, 각 격자에 구성된 파티클 가중치의 합이 1을 초과하지 않도록 제한하며, 초과하는 경우 정규화를 수행하여 확률적 일관성을 유지한다.

의미론적 동적 격자지도의 핵심 과정 중 하나는 다양한 센서 및 딥러닝 기반 모델에서 제공하는 시맨틱 정보를 불확실성과 함께 효과적으로 융합하는 것이다. BEVFusion에서 획득한 BEV 기반 객체 인식 결과와 2DPASS에서 산출된 시맨틱 포인트 정보를 격자 셀 단위로 매핑하며, 각 센서 및 알고리즘으로부터 제공되는 라벨 정보는 서로 다른 신뢰도를 가진다. 이를 해결하기 위해 Dempster-shafer 증거 이론을 기반으로 한 융합 기법을 적용한다. Dempster-shafer 이론에서는 각 증거가 특정 가설(예: 차량, 보행자, 자전거 등)에 대해 신뢰도 $m(A)$ 를 할당한다. 두 개의 독립적인 증거 m_1 과 m_2 가 주어졌을 때, 새로운 믿음 함수는 다음과 같이 계산된다.

$$(m_1 \oplus m_2)(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B)m_2(C)} \quad (1)$$

식 (1)에서 분자는 두 증거가 동일한 가설을 지지하는 정도를 나타내며, 분모는 서로 충돌하는 증거를 제외한 정규화 과정이다. 이 과정을 통해 서로 다른 센서에서 제공한 정보가 융합되며, 상충하는 경우에도 불확실성을 유지하면서 최적의 결정을 내릴 수 있다. 최종적으로, 각 객체는 점유 확률, 비점유 확률, 동적 상태, 속도, 차량 확률, 바이크 확률, 사람 확률, 기타 확률, 불확실성을 포함하는 의미론적 동적 격자지도로 완성된다. 여기서 각 객체의 상태는 파티클 필터 기반의 확률적 모델링을 통해 보정되며, 클러스터링을 통해 객체의 중심점, 형상, 속도, 헤딩을 추출하게 된다. 이렇게 구축된 의미론적 동적 격자지도는 기존 동적 격자지도 보다 한층 발전된 모델로, 단순한 점유 상태 추정과 거동 여부에 머무르지 않고 특정 객체의 형태, 라벨, 동적 특성, 신뢰도를 종합적으로 고려하여 보다 정교한 추론이 가능하다. 본 논문에서는 의미론적 동적 격자지도에서 클러스터링을 통해 얻은 객체의 라벨 정보와 상태 정보를 기반으로 확장 칼만 필터 기반 상태 추정 알고리즘 추가적으로 결합하여 더욱 강건한 객체 추적을 구현하였다.

3. 라벨 기반 객체 추적 기법

의미론적 동적 격자지도는 객체를 차량, 바이크, 사람, 기타로 분류하고, 각 객체의 중심점, 형상, 속도, 헤딩 정

보와 함께 라벨별 확률을 제공한다. 그러나 측정 오차, 오인식, 미인식이 혼재하여 시계열적으로 불완전한 정보가 발생할 수 있다. 본 연구는 이러한 불확실성을 시간 축에서 완화하고 라벨 특유의 거동을 반영하기 위해 라벨 기반 EKF³²⁾를 중심으로 한 추적 프레임워크를 설계하였다.

Fig. 2는 제안한 파이프라인의 전체 흐름을 요약한다. 먼저, 의미론적 동적 격자지도의 측정값을 로컬 좌표계에서 전역 좌표계로 변환한다. 속도 정보가 전역 기준으로 제공되므로 위치 또한 동일 기준으로 정렬하여 위치, 속도 불일치로 인한 편향을 사전에 제거한다.

예측 단계에서는 유지 중인 Tracklet은 라벨별 모션 모델과 공분산을 적용한 라벨 기반 EKF로 예측을 진행한다. 자세한 구성은 3.1절과 3.2절에서 다룬다. 정지로 판단되거나 관측 헤딩이 트랙 헤딩과 크게 어긋날 때는 Dual prediction model을 적용해 저속, 정지 구간의 방향성 불확실성을 줄여준다. 이에 대한 설명은 3.5절에서 다룬다.

데이터 결합 단계에선 Nearest neighbor³³⁾을 사용해 예측된 Tracklet과 신규 측정치를 매칭한다. 이 과정에서 3가지 경우가 발생하며, 트랙 벡터에서 M/N 로직³⁴⁾으로 관리되며, 매칭되지 않은 측정치는 임시 Tracklet으로 등록하고 일정 횟수 매칭 시 확정된다. 지속적으로 매칭되지 않은 Tracklet은 소실로 판단되어 삭제된다. 매칭이 성립하면 EKF 갱신을 통해 상태를 보정하고, 라벨 벡터를 통해 누적 라벨 확률로 최종 클래스를 확정한다. 이 부분은 3.3절에서 상세히 다룬다.

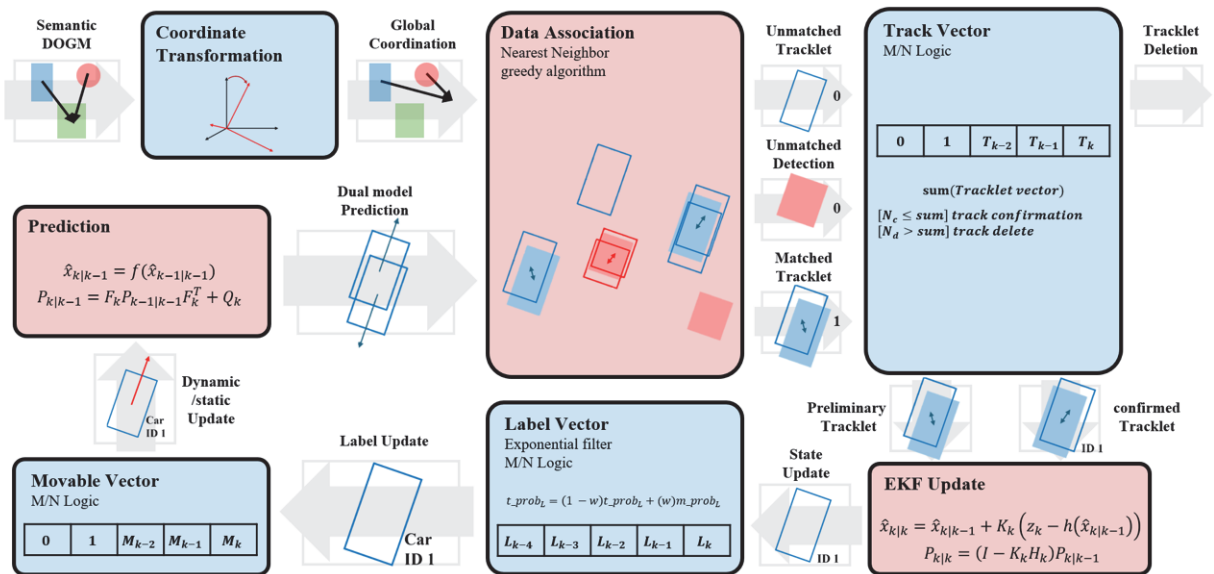


Fig. 2 Overview of the proposed multi-object tracking framework

라벨이 확정된 Tracklet에는 해당 라벨 전용 모션 모델, 공분산을 적용하여 라벨별 거동 특성을 반영한다. 또한, 무버블 벡터로 최신 속도 변화를 평가해 동/정적 상태를 구분한다. 자세한 설명은 3.4절에서 설명할 예정이다. 이때 정지로 판별된 객체는 Dual prediction model을 적용해 정지에서 저속 전이에서 발생하기 쉬운 방향성 흔들림을 최소화한다. 특히 SDOGM 환경에서는 객체가 정지 상태에서 갑작스럽게 이동하거나 큰 거동 변화를 보일 경우 파티클 기반 추정이 수렴하지 못하는 문제가 발생하는데, Dual prediction은 이를 완화하여 추적 안정성을 높이는 역할을 한다.

요약하면, 라벨별로 최적화된 모션 모델과 공분산을 적용한 EKF를 통해 예측 및 추정의 정확도를 향상시켰다. 또한 Dual prediction을 적용하여 정지, 저속 구간에서 방향성의 신뢰성을 높임으로써 매칭 성능을 개선하였다. 이러한 기여는 이후 4장에서 정량 지표를 통해 검증된다.

3.1 라벨별 모션 모델

모션 모델은 데이터 결합을 위해 이전 스텝의 Tracklet의 상태로 다음 위치를 예측하는 데 사용된다. 예측 오차가 크면 데이터 결합 신뢰도가 떨어져 전체 추적 성능이 급격히 저하되므로, 객체 라벨별로 특화된 모델을 적용했다. 차량과 바이크는 회전율과 속력을 함께 고려하는 Constant Turn Rate and Velocity(CTRV) 모델, 보행자는 선형 이동 가정의 CV모델, 기타 객체는 일정 회전만 가정한 Constant Turn(CT)모델을 사용해 각 라벨의 거동 특성을 효과적으로 반영하였다.

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ l_{k+1} \\ w_{k+1} \\ s_{k+1} \\ \psi_{k+1} \\ \dot{\psi}_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + \frac{s_k}{\dot{\psi}}(\sin(\psi_k + \dot{\psi}_k T) - \sin(\psi_k)) \\ x_k + \frac{s_k}{\dot{\psi}}(\cos(\psi_k) - \cos(\psi_k + \dot{\psi}_k T)) \\ l_k \\ w_k \\ s_k \\ \psi_k + \dot{\psi}_k T \\ \dot{\psi}_k \end{bmatrix} \quad (2)$$

식 (2)는 CTRV 모션 모델의 상태 방정식을 나타낸다. 여기서 x와 y는 위치, l과 w는 길이와 너비, s는 속도, ψ 는 헤딩, $\dot{\psi}$ 는 회전율을 의미한다. 시간 T는 샘플링 시간을 나타낸다.

사람의 경우 이동 방향의 자유로움이 높은 특징이 있다. 본 논문에서는 사람에 대해 CV 모션 모델을 적용하여, 단순하고 직관적인 이동 경로 예측을 수행하였다. 일

정한 속력을 가정하여 직선 운동을 예측하는 방식이다.

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ l_{k+1} \\ w_{k+1} \\ v_{x,k+1} \\ v_{y,k+1} \\ \psi_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + Tv_{x,k} \\ x_k + Tv_{y,k} \\ l_k \\ w_k \\ v_{x,k} \\ v_{y,k} \\ \psi_k \end{bmatrix} \quad (3)$$

식 (3)은 CV 모션 모델의 상태 방정식을 나타낸다. 여기서 x와 y는 위치, l과 w는 길이와 너비, $v_{x,k}$ 와 $v_{y,k}$ 는 각각 x 축과 y 축 방향의 속도 성분을 의미한다. 시간 T는 샘플링 시간으로, 일정한 속력을 가정하여 직선 운동을 모델링한다.

기타 객체는 사람, 차량, 바이크와 같은 특정 라벨로 명확히 분류되지 않은 객체들을 포함한다. 이러한 객체들은 고유한 움직임 패턴을 파악하기 어려우며, 일정한 직선 운동이나 회전 운동만으로는 그 거동을 정확하게 예측하기 어렵다. 본 논문에서는 이러한 불확실성을 고려하여, CT 모션 모델을 사용하였다. CT 모션 모델은 직선 운동과 회전 운동 사이의 중간 형태로, 일정한 속력과 회전을 가정하여 객체의 거동을 예측한다.

이 모델은 불확실한 객체가 직선 또는 회전 거동을 보일 가능성을 모두 반영할 수 있어, 본 연구에서는 특정 움직임 패턴을 단정 짓기 어려운 기타 라벨 객체에 적용하였다.

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ l_{k+1} \\ w_{k+1} \\ s_{k+1} \\ \psi_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + Ts_k \sin(\psi_k) \\ x_k + Ts_k \cos(\psi_k) \\ l_k \\ w_k \\ s_k \\ \psi_k \end{bmatrix} \quad (4)$$

식 (4)는 CT 모션 모델의 상태 방정식을 나타낸다. 여기서 x와 y는 위치, l과 w는 길이와 너비, s는 속도, ψ 는 헤딩을 의미한다. 시간 T는 샘플링 시간으로, 일정한 회전과 속력을 기반으로 객체의 궤적을 예측한다. 이와 같이 다양한 모션 모델을 각 라벨에 적합하게 적용함으로써, 도심지 환경에서의 객체 추적 정확성과 안정성을 효과적으로 향상시킬 수 있다.

3.2 라벨별 공분산

칼만 필터에서 공분산 행렬은 불확실성을 나타내며, 추정 상태의 분산 및 상태 변수 간의 상관관계를 포함한

다. 초기 상태 공분산 행렬 P 는 필터가 첫 번째 업데이트를 수행하기 전에 상태의 불확실성을 반영하며, 필터의 수렴 속도와 초기 추적 안정성에 영향을 미친다. 관측 노이즈 공분산 행렬 R 은 의미론적 동적 격자지도로부터 얻은 측정값의 불확실성을 모델링한다. R 은 의미론적 동적 격자지도의 라벨별 인지 성능에 따라 다르게 설정되며, 이는 측정값의 신뢰성과 추적 정확도에 직접적인 영향을 준다. 프로세스 노이즈 공분산 행렬 Q 는 시스템의 동적 모델에서 발생하는 불확실성을 반영한다. 객체의 라벨에 따라 적용되는 모션 모델이 다르므로, 프로세스 노이즈도 객체 특성에 맞게 조정된다. 예를 들어, 차량과 바이크 모두 CTRV 모션 모델을 사용하지만, 바이크는 더 민첩한 움직임을 보이므로 바이크의 Q 의 속력 및 회전율의 분산을 크게 설정하여 급격한 움직임 변화를 민감하게 반영할 수 있도록 한다. 반면, 차량은 더 큰 질량과 관성을 가지므로 차량의 Q 에서 속력 및 회전율의 분산을 비교적 작게 설정하여 안정적인 추적을 유지한다. 이와 같이 동일한 모션 모델을 사용하더라도 공분산 행렬을 라벨에 특성에 맞게 최적화하여, 객체의 동적 특성을 보다 정확하게 반영할 수 있다.

3.3 라벨 확률 관리

의미론적 동적 격자지도는 각 객체에 대해 차량, 바이크, 사람, 기타에 대한 확률 정보를 제공한다. 그러나 실제 객체가 바이크인 경우에도 항상 바이크의 확률이 가장 높게 출력되는 것은 아니며, 이로 인해 객체 라벨이 빈번히 변동하는 현상이 발생할 수 있다. 라벨이 자주 변경되면 그에 따라 모션 모델도 반복적으로 바뀌게 되어, Tracklet의 신뢰성이 저하되고 잘못된 상태 추정이 이루어질 가능성이 높아진다. 이러한 문제를 방지하기 위해, 본 논문에는 가중치 필터와 M/N 로직을 결합하여 라벨 확률을 누적하고, 이를 기반으로 Tracklet의 최종 라벨을 안정적으로 결정하였다. 해당 라벨 보정 과정은 식(5)~(8)을 통해 수식적으로 표현된다. 먼저, 측정된 각 라벨 확률 m_L 과 이전 프레임의 추적 확률 t_L 을 가중 평균하여 지속적으로 누적 확률 t_L 을 갱신한다.

$$t_{C,k} = (1-w) \cdot t_{C,k-1} + w \cdot m_{C,k} \quad (5)$$

$$t_{B,k} = (1-w) \cdot t_{B,k-1} + w \cdot m_{B,k} \quad (6)$$

$$t_{P,k} = (1-w) \cdot t_{P,k-1} + w \cdot m_{P,k} \quad (7)$$

$$t_{O,k} = (1-w) \cdot t_{O,k-1} + w \cdot m_{O,k} \quad (8)$$

여기서 w 는 가중치인 0과 1 사이의 값이다. m_L 은 현재 프레임 k 에서 측정된 라벨 L 의 확률(차량: C, 바이크: B, 사람: P, 기타: O), t_L 은 이전 프레임에서의 추적 확률이다. 식 (9)는 업데이트된 확률 t_L 중 가장 높은 값을 갖는 라벨을 선택한다.

$$L_{temp} = \operatorname{argmax}_{L \in \{C,B,P,O\}} t_{L,k} \quad (9)$$

그 후, 최근 프레임의 라벨들을 벡터 LV로 저장하여 식 (10)으로 표현할 수 있다.

$$LV = [L_{k-4}, L_{k-3}, L_{k-2}, L_{k-1}, L_{temp}] \quad (10)$$

여기서 L_{k-i} 는 과거 i 번째 프레임에서 선택된 라벨이다. 마지막으로, 라벨 벡터 LV에서 가장 많이 등장하는 라벨을 최종 확정 라벨로 지정한다. 이러한 방법을 통해 측정치와 매칭된 Tracklet에 포함된 라벨 확률들을 가중 평균하고, 신뢰성 높은 객체 히스토리를 구축하게 된다. 이 벡터를 기반으로 최종 확정 라벨을 선택함으로써 객체 라벨의 추적 일관성과 신뢰성을 높인다.

3.4 동·정적 상태 판정 및 확정 기법

의미론적 동적 격자지도에서는 파티클들의 방향 수렴을 통해 동/정적 상태를 판단한다. 그러나 객체의 급격한 거동 변화나 알고리즘 초기 단계에서 파티클들이 충분히 수렴되지 못한 경우, 잘못된 거동상태를 판단하는 현상이 발생한다. 이러한 문제를 개선하기 위해 본 논문에서는 EKF에서 추정된 속력을 기반으로 객체의 동/정적 상태를 M/N 로직을 통해 재판단하는 방법을 제안한다. 업데이트된 속력을 기반으로 객체의 상태 st_k 를 식 (11) 같이 결정한다. 이때 n 의 경우 라벨에 따라 값이 달라진다.

$$st_k = \begin{cases} 1, & \text{if } S > n [m/s] \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

결정된 상태는 M/N 로직에 저장되며, 이는 식 (12)로 표현할 수 있다.

$$MV = [st_{k-4}, st_{k-3}, st_{k-2}, st_{k-1}, st_k] \quad (12)$$

M/N로직의 총합이 3 이상일 경우 객체를 동적으로 판단하며, 그렇지 않을 경우 정적으로 판단한다. 또한, M/N

로직을 [1, 0, 0, 0, 0]으로 초기화하여 처음 관측된 객체가 동적일 경우에도 확정 트랙으로 격상된 이후 일시적으로 정적으로 판정될 수 있는 문제가 발생할 수 있다. 이러한 문제를 예방하기 위해 본 논문에서는 M/N로직을 [1, 0, 1, 1, 0]으로 초기화하여, 처음 관측된 객체가 확정 트랙 판정 전에 안정적으로 동적인 객체로 판정될 수 있도록 하였다. 이러한 방법을 통해 객체의 이동여부를 신뢰성 있게 판단할 수 있으며, 동적 객체와 정적 객체를 효과적으로 구분할 수 있다.

3.5 정적 객체를 위한 이중 예측 모델

의미론적 동적 격자지도에서 제공되는 상태값에는 속력과 헤딩 정보가 포함되어 있다. 이러한 상태 값에서 속력은 항상 양수이기 때문에 헤딩의 방향이 잘못 된다면 추적하는 과정에서 불안정한 상태값을 추정하게 된다. 그리고 도심지에는 다수의 정적 객체가 존재하며, 이들은 전진 및 후진 거동이 가능하다. 이러한 객체가 잘못 인식되거나 상태 추정이 안정적으로 추정되지 않으면 자율주행 시스템의 경로 계획과 안전에 심각한 영향을 미칠 수 있다. 또한, 딥러닝 알고리즘은 객체의 헤딩을 차량 전방으로 출력하는 경향이 있어 실제 거동방향과 불일치하는 경우가 빈번하게 발생할 수 있다. 이를 해결하기 위해 본 논문에서는 차량과 마이크 라벨에 대하여 Dual prediction model을 제안한다.

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ l_{k+1} \\ w_{k+1} \\ s_{k+1} \\ \psi_{1,k+1} \\ \dot{\psi}_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + \frac{s_k}{\dot{\psi}} (\sin(\psi_k + \dot{\psi}_k T) - \sin(\psi_k)) \\ x_k + \frac{s_k}{\dot{\psi}} (\cos(\psi_k) - \cos(\psi_k + \dot{\psi}_k T)) \\ l_k \\ w_k \\ s_k \\ \psi_{1,k} + \dot{\psi}_k T \\ \dot{\psi}_k \end{bmatrix} \quad (13)$$

식 (13)은 전진 모델의 상태 방정식을 나타낸다. 이 모델은 CTRV 모션 모델을 기반으로 하며, 인지 알고리즘에서 추정된 헤딩 값을 기준으로 한다. 여기서 x, y는 차량의 위치, l, w는 각 차량의 길이와 폭, s는 속도, ψ_1 는 인지된 객체의 진행 방향, $\dot{\psi}$ 는 방향 각속도, T는 샘플링 시간을 의미한다.

식 (14)은 후진 모델의 상태 방정식을 나타내며, 역시 CTRV 모션 모델을 기반으로 한다. 여기서 ψ_2 는 ψ_1 에 π 를 더한 값으로, 반대 방향의 헤딩을 의미한다. 이중 예측 모델 알고리즘은 정지된 객체로 판단된 경우 또는 Tracklet에 매칭된 측정값의 헤딩 방향과 예측된 방향 간

오차가 클 때 적용된다. 이를 통해 실제 객체의 이동 방향을 정확히 판단할 수 있다.

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ l_{k+1} \\ w_{k+1} \\ s_{k+1} \\ \psi_{2,k+1} \\ \dot{\psi}_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + \frac{s_k}{\dot{\psi}} (\sin(\psi_k + \dot{\psi}_k T) - \sin(\psi_k)) \\ x_k + \frac{s_k}{\dot{\psi}} (\cos(\psi_k) - \cos(\psi_k + \dot{\psi}_k T)) \\ l_k \\ w_k \\ s_k \\ \psi_{2,k} + \dot{\psi}_k T \\ \dot{\psi}_k \end{bmatrix} \quad (14)$$

Fig. 3은 이중 예측 모델 알고리즘에서 실제 거동 방향을 판단하는 과정을 보여준다. 붉은색은 측정값, 초록색은 전진 방향 모델의 예측값, 분홍색은 후진 방향 모델의 예측값을 나타낸다. 측정값과 이중 예측 모델 간의 Intersection over Union(IoU)를 통해 실제 거동 방향을 결정하며, Fig. 3처럼 IoU 값이 높은 모델이 실제 방향으로 선택된다. 이를 통해 정지 상태에서 출발하거나 저속으로 전이하는 구간에서 방향성 불안정을 줄이고, 매칭 실패 및 ID 전환을 감소시키는 효과를 얻을 수 있다.

Fig. 4는 Dual prediction model이 적용되었을 때, ID 전환이 줄고 헤딩 안정성이 개선되는 효과를 보여준다. (a)는 객체가 왼쪽에서 오른쪽으로 후진하며 가속하는 시나리오에서 인지된 객체의 이동 궤적을 나타낸다. Fig. 4의 (b)는 (a)에서 인지된 측정값을 기반으로 단일 모델 확장 칼만 필터를 수행한 결과의 궤적으로, 객체의 움직임이 불안정하게 추정된 모습을 보여준다. 색상은 ID를 나타내며, 속도 증가 후 다수의 예비 트랙과 확정 트랙이 생성되었다. Fig. 4의 (c)는 (a)에서 인지된 측정값을 기반으로 이중 예측 모델을 수행한 결과의 궤적이다. 객체 상태가 안정적으로 추정되며 하나의 ID로 일관되게 유지되는 것을 확인할 수 있다. Fig. 4의 (d)는 (a)-(c)의 객체 헤딩 값을 참값과 비교한 그래프이다. 측정값의 경우 후

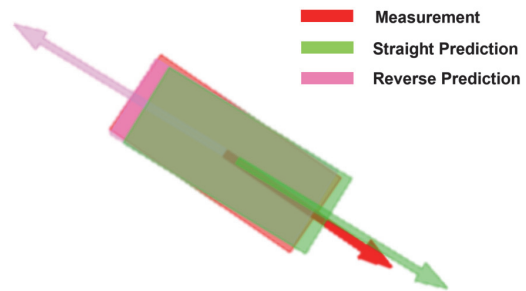


Fig. 3 Data association process using dual model

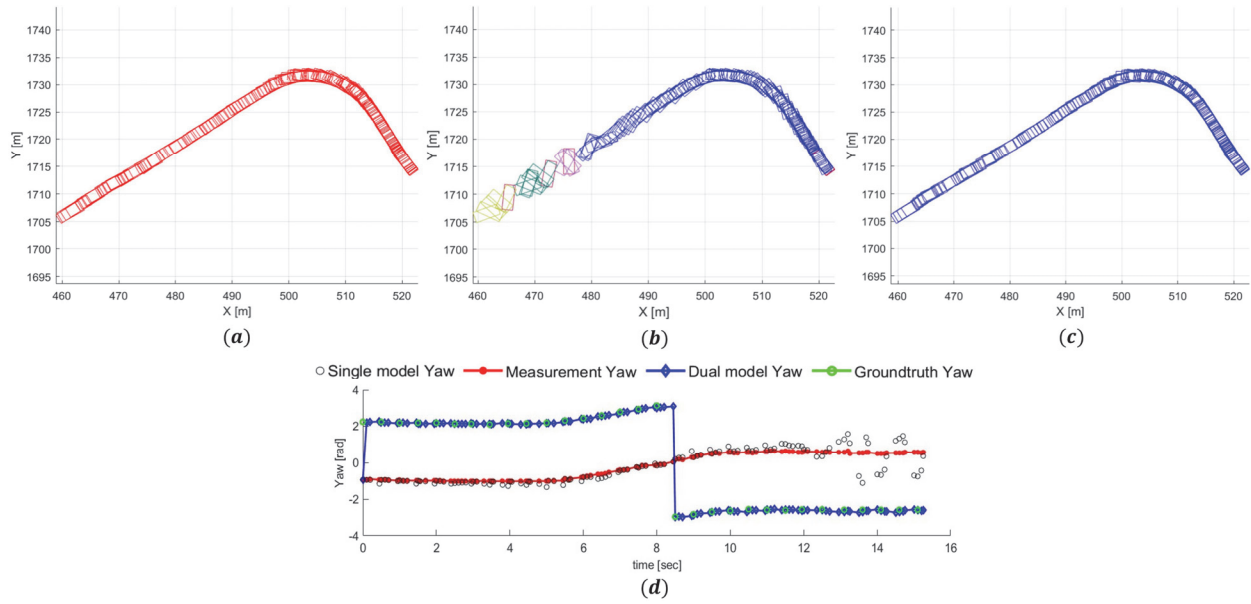


Fig. 4 Tracking results in reverse acceleration scenario: (a) measurement-based accumulated trajectory, (b) single-model Kalman filter trajectory, (c) dual-model kalman filter trajectory, (d) yaw estimation comparison among ground truth, measurement, single model, and dual model

진을 하고 있어 Ground truth와 heading이 반대값으로 나타난 것을 확인할 수 있다. Single model yaw는 단일 모델로 추정된 heading 값으로, 측정값의 heading 경향성을 따라 불안정한 추정을 보여준다. Dual model yaw는 이중 예측 모델로 추정된 heading 값으로, 실제 이동 방향을 안정적으로 추정하는 것을 확인할 수 있다. 이를 통해 제안한 이중 예측 모델은 정지·저속 또는 후진 가속과 같은 복잡한 시나리오에서 추적 안정성을 향상시키고, 다중 트랙 생성 및 ID 전환을 줄이는 효과가 있음을 알 수 있다.

4. 실험 및 결과

본 연구에서는 제안한 딥러닝 기반 의미론적 동적 격자 지도를 활용한 다중 객체 추적 알고리즘의 성능을 검증하기 위해, 총 두 개의 데이터셋과 네 개의 시나리오를 사용하였다. 시나리오는 고속 주행 환경과 복잡한 도심 환경에서의 다중 객체 추적 정확도를 평가할 수 있도록 구성되었다. 사용된 데이터셋은 공공 데이터셋인 nuScenes와 연구실의 자율주행 플랫폼을 통해 수집한 실험 데이터

셋으로 구성된다. 데이터는 자동차 주변 50 m 반경 내 객체들을 대상으로 20 Hz 주기로 전문업체에 의해 라벨링되었다.

Scene 98은 nuScenes 데이터셋 기반의 고속 주행 시나리오이다. 자동차는 직선 도로를 따라 고속으로 주행하며, 좌측 차선에는 주행 차량들이, 우측에는 정차된 차량들이 위치한다. 반대 차선에서도 다수의 차량이 이동 중이며, 전체적으로 정적 객체와 동적 객체가 혼재된 도로 구조이다. Fig. 5의 (a)는 이러한 도로 환경을 시각적으로 보여준다. Table 1에 제시된 바와 같이, 해당 시나리오에는 차량 37대, 마이크 2대, 보행자 11명이 포함되어 있다. 이 시나리오는 고속 이동 중에도 다양한 동적 및 정적 객체의 상태를 판별하고, 높은 속력을 가진 객체 간의 상태 분리 추정 정확도를 평가하는 데 적합하다.

Scene 272는 nuScenes의 도심 환경 시나리오이다. 자동차는 초기 정지 상태에서 출발하며, 동일 차선 내 전방 및 후방 차량 역시 정지 상태에서 출발한다. 주변에는 차량, 마이크, 보행자가 동시에 존재하고, 동적 상태 변화

Table 1 Label-wise object count in each scenario

Dataset	Scenario	Car	Bike	Person	Other
nuScenes	Scene 98	37	2	11	0
	Scene 272	6	3	10	0
Laboratory	Naebusoonhwan	28	0	0	0
	Sangam district	34	13	30	1

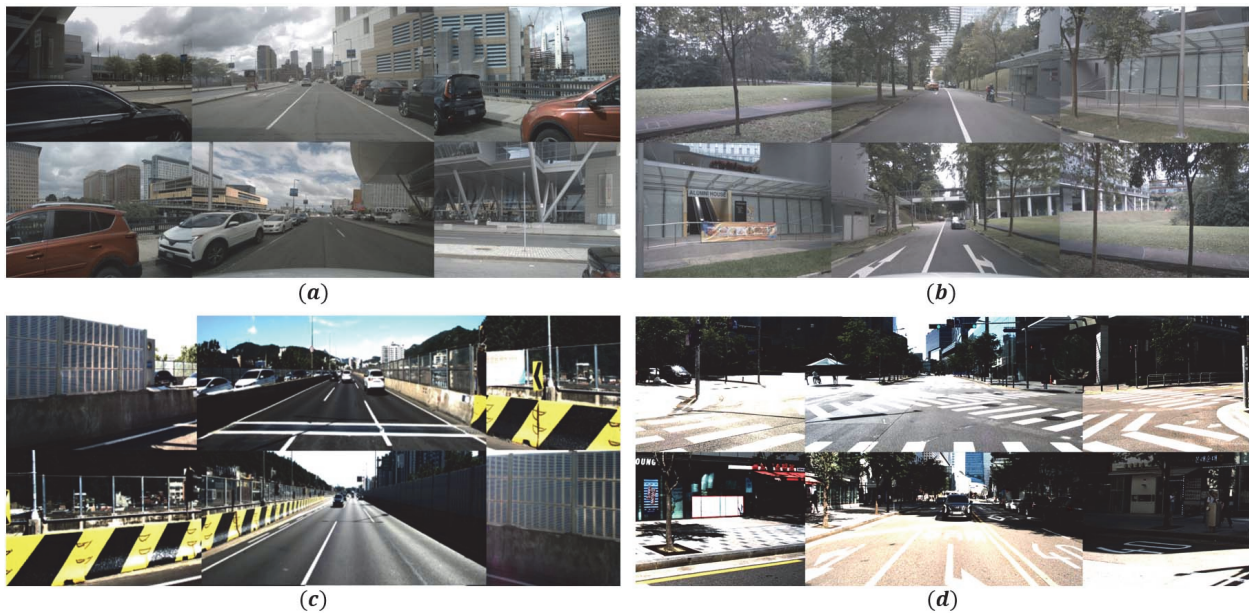


Fig. 5 Driving environments of scenarios: (a) nuScenes highway scenario Scene 98, (b) nuScenes urban scenario Scene 272, (c) highway scenario on Korea's Naebusoonhwan Road, (d) urban scenario in Korea's Sangam district

가 빈번하게 발생하는 특징을 가진다. Fig. 5의 (b)는 이러한 도심 환경을 나타내며, Table 1에 제시된 바와 같이, 해당 시나리오는 차량 6대, 바이크 3대, 보행자 10명이 포함된 복잡한 객체 구성을 보여준다. 이 시나리오를 통해 객체의 동/정적 상태를 분류하고 다수의 객체의 라벨 분류 성능을 평가하는 데 적합합니다.

내부순환도로 시나리오는 연구실 자율주행 플랫폼으로 수집된 실험용 데이터셋을 기반으로 구성되었다. 자차량은 약 75 km/h의 속력으로 주행하며, 다수의 차량이 함께 이동하는 고속 주행 환경이다. 주행 차량 외에도 감속 및 정지 구간이 포함되어 있어 속도 추정 성능과 상태 추정 정확도를 확인할 수 있다. Fig. 5의 (c)는 실제 고속 주행 장면을 보여주며, Table 1에 따르면, 이 시나리오는 차량 객체 28대로만 구성되어 단일 객체 환경을 구성한다. 객체 종류는 제한적이나 속도 기반 상태 추적 성능 평가에는 적합하다.

상암동 시나리오는 복잡한 도심 환경을 반영한다. 자차량 주변에는 정지 차량과 주행 차량이 혼재하며, 다수의 보행자는 인도를 따라 이동하고, 바이크는 도로를 교차하며 주행한다. Fig. 5의 (d)는 다양한 객체가 공존하는 도심 환경을 보여주며, Table 1에 나타난 바와 같이 이 시나리오에는 차량 34대, 바이크 13대, 보행자 30명, 기타 객체 1개가 포함되어 있어 복잡도가 높다. 이 시나리오는 객체 간 상호작용이 복잡하고 라벨 간 분포가 밀집된 상황에서, 다중 클래스 객체를 장기적으로 안정적으로 추적할 수 있는지 검증하기 위한 환경으로 적합하다.

본 연구에서 활용된 BEVFusion과 2DPASS 모델은 공개된 원저자의 학습 프로토콜을 따랐으며, NVIDIA RTX 3090 GPU 2대를 사용하여 학습을 진행하였다. 배치 사이즈의 경우 GPU 사양에 맞게 조절하였으며, 2DPASS의 경우에도 nuScenes을 사용가능하게 변환하여 학습을 진행하였다. 한국 실도로에 대한 추가 학습 없이 적용되었다.

4.1절에서는 본 논문이 제안한 기법이 실제 다양한 환경에서도 유효하게 작동하는지를 확인하기 위해, 시나리오 마다 객체의 동/정적 상태 추정, 라벨 구분 성능, 다중 객체 추적의 안정성을 정성적으로 분석하였다. 4.2절에서는 동일한 조건에서 다른 추적 기법과 객체 속도 오차와 연산시간으로 성능 비교를 진행하였다.

4.1 시나리오별 추적 성능 분석

본 절에서는 nuScenes 데이터셋과 연구실 자율주행 플랫폼을 통해 취득한 데이터셋을 활용하여, 제안한 딥러닝 기반 의미론적 동적 격자지도 기반 다중 객체 추적 알고리즘의 성능을 정량적으로 분석하였다. 각 시나리오별로 인지 단계와 추적 단계의 성능을 모두 제시하여, 제안 기법이 인지 결과를 기반으로 추적 성능을 어떻게 향상시키는지를 비교하였다. 평가 지표는 동/정적 상태 분류 정확도, 라벨 분류 정확도, 그리고 동일 객체에 할당된 추적 ID 변경 횟수이며, IDS 값이 낮을수록 추적의 연속성과 안정성이 높음을 의미한다.

Table 2는 제안 알고리즘이 모든 시나리오에서 의미론적 동적 격자지도 대비 다양한 지표에서 개선된 것을 보

Table 2 Scenario-wise evaluation of object motion state classification, label accuracy, and ID switch

Scenario	Dynamic/Static accuracy (SDOGM) ↑	Label accuracy (SDOGM) ↑	Dynamic/Static accuracy (Tracking) ↑	Label accuracy (Tracking) ↑	IDS ↓
Scene 98	90.1 %	98.4 %	93.2 %	100 %	15
Scene 272	92.8 %	98.8 %	96.8 %	100 %	2
Naebusoonhwan	97.7 %	100 %	100 %	100 %	1
Sangam district	89.1 %	97.9 %	95.3 %	100 %	23

여준다. Dynamic/Static accuracy는 객체가 주행 중인 정지 상태인지를 이분법적으로 구분한 정확도를 나타내며, Label accuracy는 차량, 보행자, 바이크 등의 객체 라벨을 얼마나 정확하게 분류하였는지를 의미한다. IDS는 동일 객체에 부여된 추적 ID가 시간 경과에 따라 잘못 변경된 횟수를 나타내며, 이는 객체 추적의 연속성과 트랙 관리 성능을 평가하는 지표로 낮을수록 안정적인 추적이 이루어졌음을 의미한다.

Scene 98은 고속 주행 중 다수의 차량과 정차 객체가 혼재된 복잡한 환경을 반영한다. 의미론적 동적 격자지도 단계에서 동/정적 정확도 90.1 %, 라벨 정확도 98.4 %를 기록하였으며, 추적 알고리즘 적용 후 각각 93.2 %와 100 %로 향상되었다. 이는 고속 이동 중에도 객체의 동/정적 상태를 효과적으로 구분했음을 보여준다. 그러나 ID 전환이 15회 발생했는데, 이는 정차 차량이 밀집된 도로 우측 구간에서 클러스터링 오류가 발생하고, 정지된 Tracklet이 전방 차량과 병합되는 과정에서 잘못된 동적 분류가 발생하였다.

Scene 272는 다양한 차량, 바이크, 보행자가 혼재하고 다수의 상태 전이가 발생하는 도심지 시나리오이다. 의미론적 동적 격자지도 단계에서 동/정적 정확도 92.8 %, 라벨 정확도 98.8 %를 기록하였으며, 추적 알고리즘 적용 후 각각 96.8 %, 100 %로 개선되었고, ID 전환은 2회에 불과했다. 이는 제안한 알고리즘이 복잡한 객체 간 상호작용과 빈번한 상태 변화에 효과적으로 대응함을 의미하며, 특히 이중 예측 모델을 통해 정지상태에서 출발하는 객체를 안정적으로 추적했기 때문이다. ID 전환은 주로 오클루전 구간에서 차량이 장기간 미인지 상태로 존재하다가 재식별되는 과정에서 발생했다.

내부순환도로 시나리오는 차량 객체만 존재하는 단일 클래스 고속 주행 환경이다. 의미론적 동적 격자지도 단계에서 동/정적 정확도 97.7 %, 라벨 정확도 100 %를 기록하였으며, 추적 알고리즘 적용 후 두 지표 모두 100 %를 유지했다. ID 전환은 단 1회 발생했으며, 이는 다양한 속력 변화 조건에서도 알고리즘의 강건성을 보여준다.

상암 시나리오는 차량, 바이크, 보행자가 고밀도로 혼

재된 복잡한 도심 환경을 반영한다. 의미론적 동적 격자지도 단계에서 동/정적 정확도 89.1 %, 라벨 정확도 97.9 %를 기록하였으며, 추적 알고리즘 적용 후 각각 95.3 %, 100 %로 향상되었다. 그러나 ID 전환은 23회로 비교적 높게 나타났으며, 이는 보행자 객체가 밀집된 횡단보도 및 인도 구간에서 클러스터링 오류와 객체 간 상호작용으로 인한 ID 재할당이 빈번하게 발생했기 때문이다.

종합적으로, 제안 기법은 의미론적 동적 격자지도 대비 모든 시나리오에서 분류 정확도를 향상시켰다. 라벨 분류 정확도는 전 시나리오에서 100 %를 유지했으며, 동/정적 상태 분류 정확도 또한 대부분 95 % 이상을 기록하였다. 이는 제안한 알고리즘이 객체의 의미 정보와 상태 변화를 효과적으로 통합하여 추적 성능을 향상시킬 수 있음을 보여준다. 다만, 객체 밀집도가 높고 상호작용이 빈번한 환경에서는 클러스터링 오류로 인해 ID 전환 빈도가 증가하는 경향이 나타났다.

4.2 추적 성능 비교 분석

본 절에서는 제안하는 추적 기법의 속력 추정 정확도와 연산 효율성을 검증하기 위해, 라벨 정보를 사용하지 않는 추적 기법과 성능을 비교하였다. 비교 대상으로 선정한 Single model 구조는 본 논문에서 제안한 알고리즘에서 라벨 기반 요소를 제거하여, 차량 모션 모델과 해당 공분산만을 적용한다. 또한, 라벨 관리 등 부가적인 안정화 기법은 포함하지 않는다. 이러한 구조는 다양한 객체가 혼재된 환경에서도 동일한 모션 모델을 적용하므로, 라벨 정보 활용 여부에 따른 성능 차이를 분석하기 위한 비교 대상으로 적합하다.

성능 평가는 절대 평균 속력 오차와 평균 연산시간 지표를 사용하였다. 절대 평균 속력 오차는 추적 상태 정확도를 나타내며, 평균 연산시간은 프레임 단위 처리 효율성을 나타낸다.

Table 3은 일반적인 하나의 모델만 사용한 칼만필터와 속력 추정 성능 비교 결과이다. 두 알고리즘은 측정치인 의미론적 동적 격자지도 보다 개선되었으며, 제안 기법은 모든 시나리오에서 Single model 대비 낮은 속력 오차

Table 3 Performance comparison of absolute mean speed error for different multi-object tracking algorithms

Method	Scene 98	Scene 272	Naebusoonhwan	Sangam district
SDOGM	0.634 m/s	0.553 m/s	0.524 m/s	0.545 m/s
Single model	0.609 m/s	0.503 m/s	0.432 m/s	0.605 m/s
Ours	0.518 m/s	0.412 m/s	0.432 m/s	0.420 m/s

Table 4 Performance comparison of average processing time for different multi-object tracking algorithms

Method	Scene 98	Scene 272	Naebusoonhwan	Sangam district
Single model	0.283 ms	0.137 ms	0.129 ms	0.357 ms
Ours	0.304 ms	0.142 ms	0.144 ms	0.374 ms

를 기록하였다. Scene 98에서는 0.518 m/s로, Single model의 성능인 0.609 m/s 대비 약 15 % 감소하였고, Scene 272에서는 0.412 m/s로 0.503 m/s 대비 약 18 % 개선되었다. 특히 상암에서는 0.420 m/s로, Single model의 성능인 0.605 m/s 대비 약 30 % 향상되었는데, Table 1에서 확인할 수 있듯이 해당 시나리오는 다양한 라벨이 혼재되어 있어 라벨별 모션 모델 적용의 효과가 크게 나타났다.

반면, 내부순환 시나리오에서는 두 기법 모두 0.432 m/s로 동일한 결과를 보였다. 이는 Table 1에서 보듯이 차량 라벨만 존재하는 환경이므로, Single model 역시 제안 기법과 동일한 차량 모션 모델과 공분산을 적용했기 때문이다.

이러한 결과는 라벨별 최적화된 모션 모델과 공분산을 적용하는 제안 기법이, 특히 다양한 객체 라벨이 존재하는 복잡한 환경에서 단일 모델 대비 속력 추정 정확도를 유의미하게 향상시킬 수 있음을 보여준다.

Table 4는 각 알고리즘의 평균 연산시간 비교 결과 나타낸다. 제안 기법은 모든 시나리오에서 Single model과 유사한 수준의 처리 속도를 유지하였다. Scene 98에서는 0.304 ms로 Single model의 성능인 0.283 ms 대비 약 0.021 ms 차이에 불과하며, Scene 272와 내부순환에서도 각각 0.005 ms, 0.015 ms의 차이만 나타났다. 상암과 같이 객체 수가 많고 라벨 구성이 복잡한 시나리오에서도 0.374 ms로, Single model의 성능인 0.357 ms와 근접한 결과를 기록하였다.

제안 기법의 연산시간이 Single model보다 소폭 증가한 이유는, 라벨별 모션 모델 적용을 위한 라벨 관리기 절차가 포함되었기 때문이다. 평균적으로 0.0145 ms 연산량이 증가하였으나 이는 실시간 처리에 무시할 수 있는 수준이다.

종합적으로, 제안 기법은 라벨 관리기와 라벨별 모션 모델 적용을 통해 다양한 라벨이 혼재된 환경에서 Single model 대비 속력 추정 오차를 최대 약 30 %까지 감소시

켰다. 이러한 정확도 향상에도 불구하고, 단일 필터 기반 구조를 유지하여 추가 연산이 전체 처리 속도에 미치는 영향을 최소화하였다. 그 결과, 평균 연산시간은 모든 시나리오에서 Single model과 유사한 수준을 유지하였으며, 이는 제안 기법이 실시간 처리가 요구되는 자율주행 환경에서 정확성과 효율성을 동시에 만족시킬 수 있음을 보여준다.

5. 결론

본 연구에서는 도심 환경에서의 객체 추적 강건성을 향상시키기 위해, 딥러닝 기반 의미론적 동적 격자지도를 활용한 객체 라벨 기반 추적 기법을 제안하였다. 인지 과정에서 발생할 수 있는 미인식, 오인식 및 측정 노이즈를 보정하기 위해 확장 칼만 필터와 M/N 로직을 결합하였으며, 라벨에 따라 최적의 모션 모델과 공분산을 지정하였다. 또한, 정지 객체의 이동 방향을 정확하게 예측하기 위해 Dual prediction model을 도입하여, 단일 예측 모델 기반 방식보다 안정적이고 신뢰성 있는 추적 성능을 확보하였다.

제안 기법은 nuScenes 데이터셋과 연구실 자율주행 플랫폼을 통해 수집한 데이터셋을 기반으로 검증하였다. 실험 결과, 다양한 도심 및 고속 주행 환경에서 속력 오차 감소와 동·정적 상태 판단 정확도 향상 등 전반적인 성능 개선이 확인되었다. 특히, 라벨 정보를 사용하지 않는 Single model과의 비교에서, 제안 기법은 라벨 관리기와 라벨별 모션 모델 적용을 통해 속력 추정 오차를 최대 약 30 %까지 감소시키면서도, 모든 시나리오에서 연산 시간을 Single model과 유사한 수준으로 유지하였다. 이는 단일 필터 기반 구조를 유지하여 불필요한 연산을 억제하는 동시에, 라벨별 최적화된 모션 모델과 공분산 적용을 통해 다양한 객체 라벨이 혼재된 환경에서도 추적 성능을 향상시킨 결과이다.

본 연구는 객체 추적 성능 향상에 기여하며, 향후 다양

한 센서 데이터를 융합하여 추적 정보의 신뢰도를 높이고, 딥러닝 기반 기법을 활용해 객체의 형상 및 거동을 정밀하게 예측함으로써, 보다 안정적이고 정확한 추적 성능을 달성할 수 있을 것으로 기대된다. 이를 통해 다양한 객체에 대한 지속적이고 일관된 추적을 실현하여, 도심지 자율주행 환경에서의 추적 안정성과 신뢰성을 한층 강화할 수 있을 것으로 전망된다.

후 기

2025년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0020536, 2025년 산업혁신인재성장지원사업). 또한, 본 연구는 2025년도 산업통상자원부 및 한국전자기술연구원(KEIT) 연구비 지원에 의한 연구임(RS-2024-00445826).

References

- 1) Seoul Autonomous Driving Vision 2030, Seoul Metropolitan Government, <https://english.seoul.go.kr/policy/transportation/seoul-autonomous-driving-vision-2030/>, 2025.
- 2) Waymo One: The World's First Fully Autonomous Ride-Hailing Service, Waymo LLC, <https://waymo.com/waymo-one/>, 2025.
- 3) Y. O. Lee and Y. S. Son, "Competitiveness Analysis of Integrated Minimal Risk Maneuver of Autonomous Vehicle Based on Domestic Patent Data," Transactions of KSAE, Vol.30, No.6, pp.477-483, 2022.
- 4) D. Bogdoll, M. Nitsche and J. M. Zöllner, "Anomaly Detection in Autonomous Driving: A Survey," arXiv preprint arXiv:2204.07974, 2022.
- 5) S. Meng, D. Shao, J. Guo and S. Gao, "Tracking without Label: Unsupervised Multiple Object Tracking via Contrastive Similarity Learning," Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023.
- 6) Z. Lu, B. Shuai, Y. Chen, Z. Xu and D. Modolo, "Self-Supervised Multi-Object Tracking with Path Consistency," arXiv preprint arXiv:2404.05136, 2024.
- 7) Z. Guan, Z. Wang, G. Zhang, L. Li, M. Zhang, Z. Shi and N. Jiang, "Multi-Object Tracking Review: Retrospective and Emerging Trend," Artificial Intelligence Review, Vol.58, Paper No.235, 2025.
- 8) G. Ciaparrone, F. L. Sánchez, S. Tabik, L. Troiano, R. Tagliaferri and F. Herrera, "Deep Learning in Video Multi-Object Tracking: A Survey," Neurocomputing, Vol.381, pp.61-88, 2020.
- 9) M. R. Adzemovic, "Deep Learning-Based Multi-Object Tracking: A Comprehensive Survey from Foundations to State-of-the-Art," arXiv preprint arXiv:2506.13457, 2025.
- 10) H. Jang, T. Kim, K. Ahn, S. Jeon and Y. Kang, "Dynamic Occupancy Grid Map with Semantic Information Using Deep Learning-Based BEVFusion Method with Camera and LiDAR Fusion," Sensors, Vol.24, No.9, Paper No.2828, 2024.
- 11) D. C. Kim, J. H. Han and M. H. Sunwoo, "Multiple Vehicle Tracking in Urban Environment Using Integrated Probabilistic Data Association Filter with Single Laser Scanner," Transactions of KSAE, Vol.21, No.4, pp.33-42, 2013.
- 12) X. Zhou, V. Koltun and P. Krähenbühl, "Tracking Objects as Points," Proceedings of the European Conference on Computer Vision (ECCV), pp.474-490, 2020.
- 13) L. Leal-Taixé, A. Milan, I. Reid, S. Roth and K. Schindler, "MOTChallenge 2015: Towards a Benchmark for Multi-Target Tracking," arXiv preprint arXiv:1504.01942, 2015.
- 14) X. Li, T. Xie, D. Liu, J. Gao, K. Dai, Z. Jiang, L. Zhao and K. Wang, "Poly-MOT: A Polyhedral Framework for 3D Multi-Object Tracking," arXiv preprint arXiv:2307.16675, 2023.
- 15) L. Wang, X. Zhang, W. Qin, X. Li, L. Yang, Z. Li, L. Zhu, H. Wang, J. Li and H. Liu, "CAMO-MOT: Combined Appearance-Motion Optimization for 3D Multi-Object Tracking with Camera-LiDAR Fusion," arXiv preprint arXiv:2209.02540, 2022.
- 16) J. J. Yu and Y. S. Kang, "Performance Analysis on the IMM-PDAF Method for Longitudinal and Lateral Maneuver Detection Using Automotive Radar Measurements," Journal of Institute of Control, Robotics and Systems, Vol.21, No.3, pp.224-232, 2015.
- 17) N. Engel, S. Hörmann, P. Henzler and K. C. J. Dietmayer, "Deep Object Tracking on Dynamic Occupancy Grid Maps Using RNNs," Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC), pp.3852-3858, 2018.
- 18) Ö. Erkent, D. S. Gonzalez, A. Paigwar and C. Laugier, "GridTrack: Detection and Tracking of Multiple Objects in Dynamic Occupancy Grids," Proceedings of the International Conference on Vision Systems (ICVS), pp.1-14, 2021.
- 19) D. Stumper, F. Gies, S. Hörmann and K. Dietmayer, "Offline Object Extraction from Dynamic Occupancy Grid Map Sequences," Proceedings of the IEEE

- Intelligent Vehicles Symposium (IV), pp.2071-2076, 2018.
- 20) J. Kälble, S. Wirges, M. Tatarchenko and E. Ilg, "Accurate Training Data for Occupancy Map Prediction in Automated Driving Using Evidence Theory," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.5281- 5290, 2024.
 - 21) K. S. Mann, A. Tomy, A. Paigwar, A. Renzaglia and C. Laugier, "Predicting Future Occupancy Grids in Dynamic Environment with Spatio-Temporal Learning," arXiv preprint arXiv:2205.03212, 2022.
 - 22) H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan and O. Beijbom, "nuScenes: A Multimodal Dataset for Autonomous Driving," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.11618-11628, 2020.
 - 23) F. Piewak, T. Rehfeld, M. Weber and J. M. Zöllner, "Fully Convolutional Neural Networks for Dynamic Object Detection in Grid Maps," Proceedings of the IEEE Intelligent Vehicles Symposium (IV), pp.1-6, 2017.
 - 24) Y. J. Kim, G. J. Kyung, S. H. Hong and Y. S. Kang, "Clustering-Based Object Recognition Using Dynamic Occupancy Grid Map (DOGM)," KSAE Spring Conference Proceedings, pp.1033-1034, 2022.
 - 25) M. P. Ronecker, X. Diaz, M. Karner and D. Watzenig, "Deep Learning-Driven State Correction: A Hybrid Architecture for Radar-Based Dynamic Occupancy Grid Mapping," Proceedings of the IEEE Intelligent Vehicles Symposium (IV), pp.2184-2191, 2024.
 - 26) Z. Liu, H. Tang, A. Amini, X. Yang, H. Mao, D. L. Rus and S. Han, "BEVFusion: Multi-Task Multi-Sensor Fusion with Unified Bird's-Eye View Representation," Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp.8940-8947, 2023.
 - 27) X. Yan, J. Gao, C. Zheng, C. Zheng, R. Zhang, S. Cui and Z. Li, "2DPASS: 2D Priors Assisted Semantic Segmentation on LiDAR Point Clouds," Proceedings of the European Conference on Computer Vision (ECCV), pp.677-695, 2022.
 - 28) D. R. Ahn, S. G. Song, H. G. Heo and J. O. Jung, "Sensor Fusion for Pedestrian Detection Using Dempster-Shafer Theory," KSAE Fall Conference Proceedings, pp.623-627, 2015.
 - 29) Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," Proceedings of the IEEE, Vol.86, No.11, pp.2278-2324, 1998.
 - 30) H. Jang, J. Jo, S. Heo and Y. Kang, "Performance Improvement of Deep Learning Object Detection Method Using Dynamic Occupancy Grid Map," Transactions of KSAE, Vol.30, No.10, pp.839-847, 2022.
 - 31) F. Neumann, F. Deroo, G. von Wichert and D. Burschka, "Particle-Based Dynamic Semantic Occupancy Mapping Using Bayesian Generalized Kernel Inference," Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC), pp.1-6, 2024.
 - 32) D. Simon, Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches, Hoboken, John Wiley & Sons, 2006.
 - 33) P. D. Konstantinova, A. Udvarov and T. Semerdjiev, "A Study of a Target Tracking Algorithm Using Global Nearest Neighbor Approach," Proceedings of the International Conference on Computer Systems and Technologies (CompSysTech 2003), pp.290-295, 2003.
 - 34) S. S. Blackman and R. Popoli, Design and Analysis of Modern Tracking Systems, Norwood, Artech House, 1999.