클래스 균형 데이터 큐레이션(CBHEM)을 통한 자율주행 버스의 구간 선택적 모방학습 기법

오연주, 김도균, 원치훈

Class-Balanced Data Curation (CBHEM) for Segment-Selective Imitation Learning in Autonomous Buses

요 약

자율주행 버스는 고정 노선의 이점에도, 보수적 주행 정책으로 인한 급감속·저속과 잦은 개입 문제를 안는다. 본 연구는 새 주행 모델 대신, 취약 구간만 선택적으로 보강하는 데이터 큐레이션 방법론 CBHEM(Class-Balanced Hard Example Mining)을 제안한다. CBHEM은 OHEM과 Class-Balanced Learning을 결합해, 세 실패 범주(급정거·차선변경·곡선차로)의 비율을 고정한 채 통계적으로 선별한 취약 구간을 규칙 기반 expert의 정답으로 보강한다. PilotNet·CARLA를 표준 도구로 고정한 통제 비교로 데이터 전략의 기여를 분리하였다. 균형을 무시한 큐레이션이 소수 범주를 망각한 반면, CBHEM은 이를 방지하며 취약 구간 오차를 학습 전 대비 유의하게 개선하였고 반복에도 안정적이었다. 이로써 적은 데이터로 약점을 효율적으로 보강하며, 완전 무인 단계로의 전환 가능성을 제시한다.

목 차

Ⅰ. 서론

1. 연구 배경

자율주행 기술은 대중교통 사각지대 해소와 운전 인력 부족 문제의 해법으로 주목받으며 빠르게 실증 단계에 진입하고 있다. 특히 버스는 정해진 노선을 반복 운행하므로, 동일 구간을 여러 차례 학습하고 취약 지점을 반복 추적·보강할 수 있어 기술 고도화에 유리하다.

그러나 실제 도로에서의 운행은 여전히 안전관리원의 상시 동승과 개입을 전제로 한다. 경기도 판교 '판타G버스'의 운행에서 급정거, 차선 변경 망설임, 정류장 진입 시의 과도한 정지와 같은 불안정한 거동이 반복적으로 관측되었으며, 그때마다 안전관리원이 제어에 개입하였다. 이러한 현상의 근본 원인은 자율주행 시스템의 안전 중심적·보수적 주행 정책에 있다. 보수적 정책은 충돌을 회피하는 데에는 유리하지만, 정상적인 흐름에서도 과도하게 감속·정지하여 승차감을 떨어뜨리고 잦은 개입을 유발한다.

2. 문제 정의

기존 자율주행 시스템의 한계는 두 가지로 정리된다. 첫째, 시스템이 어느 구간에서 취약한지를 사후적으로 파악하더라도 이를 학습에 체계적으로 반영하는 폐루프(closed-loop) 구조가 부재하다. 안전관리원의 개입은 매일 발생하지만, 그 개입이 가리키는 실패 지점의 정보가 다시 모델 개선으로 환류되지 못한 채 일회성 기록으로만 남는 경우가 많다. 둘째, 취약 구간의 데이터를 확보하더라도 실패 양상별 수집량이 크게 달라, 특정 양상의 데이터를 단순히 많이 모아 가중 학습하면 그 상황에는 강해지는 대신 이전에 잘 처리하던 다른 상황을 잊는 파국적 망각(catastrophic forgetting)이 발생한다.

따라서 자율주행 버스가 안전관리원 동승 단계에서 완전 무인 단계로 점진적으로 진화하기 위해서는, 개입이 가리키는 약점 지점의 데이터를 지속적으로 누적·활용하면서도 전체 주행 능력의 균형을 유지하는 학습 체계가 요구된다. 본 연구는 이 두 한계를 동시에 해소하는 데이터 큐레이션 방법을 제안한다.

3. 연구 목적 및 접근 방법

본 연구는 자율주행의 성능을 좌우하는 것이 반드시 더 정교한 모델만은 아니라는 문제의식에서 출발한다. 최근 인공지능 분야에서는 모델 구조를 개선하기보다 어떤 데이터를 어떻게 선별하여 학습시킬 것인가에 주목하는 데이터 중심(Data-Centric AI) 접근이 새로운 패러다임으로 제시되고 있다. 본 연구는 이러한 흐름을 고정 노선형 자율주행 버스 환경에 적용하여, 새로운 자율주행 모델이 아닌 취약 지점만을 선택적으로 보강하는 데이터 큐레이션 방법론을 제안한다. 이를 구현한 시스템을 VITAL(Vision Improvement via Takeover-driven Adaptive Learning)이라 한다.

본 연구의 기여는 다음 네 가지로 요약된다. 첫째, 데이터 큐레이션 기법인 CBHEM(C

lass-Balanced Hard Example Mining)을 제안한다. 이는 어려운 예시에 집중하는 OHEM과 범주 간 균형을 유지하는 Class-Balanced Learning을 결합한 것으로, 세 실패 범주(급정거·차선변경·곡선차로)의 비율을 고정한 채 약점 구간만을 보강함으로써 파국적 망각 없이 데이터 효율을 높인다. 둘째, 약점을 진단하는 역할과 정답을 제공하는 역할을 분리한다. 보수적 주행 정책의 실패는 어느 구간이 취약한지를 식별하는 진단 신호로 활용하고, 규칙 기반 expert가 해당 구간의 정답 조작값을 제공하여, 일관성이 낮은 수동 조작에 의존하지 않는 학습 구조를 설계한다. 셋째, 제안 방법의 효과를 통제 비교로 입증한다. 학습 모델인 PilotNet과 시뮬레이터인 CARLA를 표준 도구로 고정한 채 데이터 전략만을 달리하여, 절대 성능이 아니라 동일 조건에서의 상대 비교(ablation)를 통해 데이터 전략의 기여를 분리하여 확인한다. 넷째, 통계적으로 엄밀한 검증을 수행한다. 취약 구간의 개선을 학습 전 보수 정책 대비로 측정하고, 그 차이를 부트스트랩 신뢰구간으로 검정하며, 큐레이션을 반복했을 때의 성능 안정성까지 확인한다.

4. 논문의 구성

본 논문의 구성은 다음과 같다. Ⅱ장에서는 End-to-End 모방학습, 운전자 개입 기반 모델 개선, 데이터 불균형 해소 및 데이터 큐레이션 기법 등 관련 연구를 검토하고 본 연구의 차별성을 정리한다. Ⅲ장에서는 VITAL의 페루프 아키텍처와 핵심 기여인 CBHEM 데이터 큐레이션 방법론, 그리고 학습 모델 구현을 기술한다. Ⅳ장에서는 CARLA 기반 실험 환경과 데이터 수집 구성, 평가 지표 및 비교군, 통계적 유의성 검정 방법을 제시한다. Ⅴ장에서는 약점 구간 식별 결과와 데이터 전략별 통제 비교, 망각 방지와 약점 개선, 반복 사이클 동역학을 분석한다. Ⅵ장에서는 시뮬레이션 환경의 한계와 이에 대한 통제, 그리고 연구의 확장 방향을 논의하고, Ⅶ장에서 결론과 기대효과·활용방안을 제시한다.

Ⅱ. 선행 연구 분석

본 연구가 제안하는 시스템의 이론적 토대가 되는 선행 연구를 세 가지 흐름으로 나누어 검토한다. 첫째는 카메라 영상을 직접 조작값으로 변환하는 End-to-End 모방학습 연구이며, 둘째는 운전자 개입(Takeover/Disengagement) 데이터를 모델 개선에 활용하는 연구, 셋째는 학습 데이터의 불균형을 해소하고 약점만 보강하는 데이터 큐레이션 기법이다. 마지막으로 이들 연구의 한계를 정리하고 본 연구의 차별성을 제시한다.

1. End-to-End 모방학습 기반 자율주행

자율주행의 제어 방식은 인지·판단·제어를 별도 모듈로 나누는 모듈형 방식과, 카메라 입력에서 조작값까지를 하나의 신경망으로 처리하는 End-to-End 방식으로 구분된다. End-to-End 방식의 대표적 연구는 NVIDIA가 제안한 PilotNet[1]이다. PilotNet은 전방 카메라의 원시 픽셀을 입력받아 조향각을 직접 출력하는 합성곱 신경망(CNN)으로, 사람이 운전한 영상과 조작값의 쌍만으로 차선 유지와 같은 주행을 학습할 수 있음을 보

였다. 이처럼 전문가의 시연을 모방하여 입력과 조작의 대응 관계를 학습하는 방식을 행동 복제(Behavior Cloning)라 한다.

그러나 단순 행동 복제는 공변량 이동(Covariate Shift) 문제를 안고 있다. 모방학습 모델은 전문가가 주행한 상태만 학습하기 때문에, 실제 주행 중 작은 오차가 누적되어 학습 데이터에 없던 상태로 벗어나면 회복하지 못하고 오류가 점차 확대된다. 이를 완화하기 위해 Ross 등[2]은 DAgger(Dataset Aggregation)를 제안하였다. DAgger는 학습된 주행 모델을 직접 주행시키면서 모델이 방문하는 상태에 대해 전문가가 정답을 추가로 제공하고, 이를 누적 학습함으로써 분포 이동에 강건한 정책을 얻는다. 다만 DAgger는 모델이 방문하는 모든 상태에 대해 사람 전문가가 일일이 정답을 제공해야 하므로 라벨링 부담이 크다.

2. 운전자 개입 데이터 기반 주행 모델 개선

자율주행차가 도로에서 운행되면서, 안전관리원이 위험을 감지해 제어권을 넘겨받는 개입(Takeover) 또는 이탈(Disengagement) 사례가 점차 늘고 있다. 개입 사례는 주행 모델이 정상 주행에 실패한 구체적 지점을 알려주는 중요한 신호로서, 이를 학습에 활용하려는 연구가 활발히 진행되고 있다.

Zhou 등[3]이 제안한 DRARL(Disengagement-Reason-Augmented Reinforcement Learning)은 개입 데이터를 그대로 학습에 사용할 때 발생하는 문제를 지적했다. 첫째, 개입 사례는 대개 드물게 발생하여 학습에 충분한 양을 확보하기 어렵다. 둘째, 모든 개입이 주행 모델의 실패에서 비롯되는 것은 아니므로, 운전자가 무심코 개입한 우연한 사례는 걸러내야 한다. 셋째, 정책을 올바로 개선하려면 실패의 실제 원인, 즉 어떤 위험 객체나 분포 외(Out-of-Distribution) 상태가 실패를 유발했는지를 식별해야 한다. DRARL은 분포 외 상태 추정 모델로 개입의 원인을 자동으로 판별하여 우연한 개입을 배제하고, 그 원인을 보존한 가상 환경에서 강화학습으로 주행 모델을 개선한다. Liu 등[4]이 제안한 TakeAD는 사전 학습된 모방학습 모델을 개입 데이터로 미세조정하는 후처리 최적화 프레임워크로, 오픈루프 학습과 클로즈드루프 배포 사이의 불일치가 개입을 유발한다는 문제의식에서 출발하여 DAgger 기반 반복 모방학습과 선호도 최적화(Direct Preference Optimization, DPO)를 결합해 폐루프 주행 성능을 향상시켰다.

본 연구와 가장 직접적으로 비교되는 연구는 Yang 등[5]의 DTCCL(Disengagement-Triggered Contrastive Continual Learning)이다. DTCCL은 자율주행 버스를 대상으로 한다는 점에서 본 연구와 같은 문제 영역을 다루며, 버스 노선의 이탈 사례가 상호작용이 많은 특정 구간에 지리적으로 집중된다는 점에 주목한다. DTCCL은 이탈이 발생할 때마다 주변 에이전트를 교란하여 양성·음성 표본을 생성하는 데이터 증강을 수행하고, 대조 학습으로 안전한 거동과 위험한 거동을 구분하도록 정책의 표현을 정교화하며, 이를 클라우드-엣지 루프에서 지속적으로 갱신함으로써 이탈이 잦은 구간의 성능을 보강하면서도 기존 구간의 능력을 유지하고자 하였다. 이들 연구는 개입·이탈 데이터가 모델 개선의 핵심 신호임을 공통적으로 보여준다. 다만 DRARL과 TakeAD는 강화학습이나 선

호도 최적화를 도입하여 구현 복잡도가 높고, 개입 데이터를 정답으로 직접 사용하는 접근은 그 정답의 품질이 개입자의 조작 일관성에 좌우된다는 한계를 갖는다.

3. 데이터 불균형 해소와 어려운 예제 학습

개입 데이터를 활용할 때 또 다른 핵심 과제는 데이터의 불균형이다. 급정거·차선변경·곡선차로 등 실패 양상별로 수집되는 데이터의 양이 크게 달라, 특정 양상에 학습이 치우치면 다른 양상의 성능이 저하될 수 있다. Shrivastava 등[6]이 제안한 OHEM(Online Hard Example Mining)은 학습 중 손실(loss) 값이 높은 예시, 즉 모델이 어려워하는 예시만을 선별하여 집중적으로 학습에 반영하는 기법으로, 쉬운 예시에 학습이 낭비되는 것을 막고 모델의 약점을 효율적으로 보강한다. 한편 Cui 등[7]은 Class-Balanced Loss를 제안하여, 단순한 샘플 개수가 아닌 유효 샘플 수(Effective Number of Samples) 개념으로 각 클래스의 기여도를 재조정함으로써 클래스 간 불균형 문제를 완화하였다.

그러나 OHEM처럼 어려운 예시만 강조하면 특정 범주에 학습이 쏠려 기존에 잘하던 능력을 잊는 파국적 망각(Catastrophic Forgetting)이 발생할 수 있다. 본 연구는 이 두 기법의 장점을 결합하여, 범주 간 균형을 고정한 채 약점 구간을 어려운 예시로 보강하는 데이터 큐레이션 기법 CBHEM(Class-Balanced Hard Example Mining)을 제안한다. 구체적인 로직은 Ⅲ장 3절에서 상술한다.

4. 선행 연구의 한계와 본 연구의 차별성

선행 연구를 종합하면 다음과 같은 한계가 확인된다. 첫째, PilotNet[1]과 같은 순수 모방학습은 공변량 이동에 취약하고, 이를 보완하는 DAgger[2] 계열은 사람 운전자의 라벨링 부담이 크다. 둘째, 개입 데이터 기반 연구 중 DRARL[3]과 TakeAD[4]는 강화학습이나 선호도 최적화를 활용하여 구현 복잡도가 높고, 고정 노선이라는 버스의 특성을 직접 활용하지는 않는다. 셋째, 버스를 직접 대상으로 한 DTCCL[5]은 이탈 구간의 지리적 집중을 활용하지만, 에이전트 교란 증강과 대조 표현 학습으로 모델 내부를 개선하는 데 초점을 두며, 실패 범주 간 비율을 명시적으로 고정해 망각을 방지하는 데이터 구성의 문제는 다루지 않는다.

본 연구가 제안하는 VITAL은 이러한 한계를 다음과 같이 보완한다. 첫째, 버스의 고정 노선·반복 운행 특성을 활용하여 노선을 정류장 단위 구간으로 분할하고, 성능이 낮은 구간만 선택적으로 보강하는 구간 선택적 학습을 수행한다. 둘째, 강화학습의 복잡한 보상 설계 없이 PilotNet 기반 행동 복제만으로 학습하여 구현 부담을 낮춘다. 셋째, 개입과 정답의 역할을 분리한다. 즉 보수적 정책의 실패는 약점 구간을 식별하는 진단 신호로만 활용하고, 정답 조작값은 일관성이 보장된 규칙 기반 expert로부터 확보함으로써, 개입 조작의 품질에 좌우되던 기존 접근의 한계를 해소한다. 넷째, OHEM[6]과 클래스 균형 학습[7]을 결합한 CBHEM을 통해, 모델 내부가 아니라 학습 데이터의 구성 층위에서 약점을 보강하면서도 파국적 망각을 방지한다. 특히 DTCCL과 달리 본 연구의 기여는 특정 모델의 표현 학습이 아니라 데이터 구성 전략에 있으므로, 학습 모델을 교체하

더라도 동일하게 재사용할 수 있다.

Ⅲ. VITAL: 운전자 개입 기반 구간 선택적 학습

본 장에서는 제안하는 데이터 큐레이션 방법론과 이를 구현한 VITAL(Vision Improvement via Takeover-driven Adaptive Learning) 시스템을 기술한다. 1절에서 본 연구의 관점, 즉 제안 방법론과 검증 도구의 구분을 분명히 하고, 2절에서 시스템 전체를 관통하는 폐루프 아키텍처를 제시하며, 3절에서 본 연구의 핵심 기여인 CBHEM 데이터 큐레이션 방법을 설명한다. 4절에서 학습 모델의 구현을 기술한다.

1. 연구의 관점

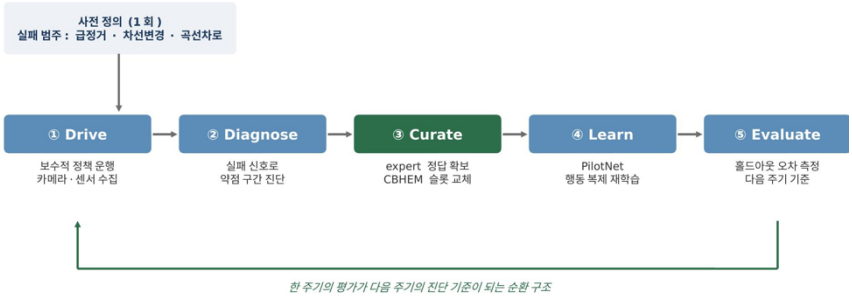
본 연구를 정확히 이해하기 위해서는 제안 대상과 검증 도구를 명확히 구분할 필요가 있다. 본 연구는 새로운 자율주행 모델을 제시하는 연구가 아니라, 방대한 데이터를 요구하는 자율주행 학습에서 취약 지점만을 선별하여 효율적으로 보강하는 데이터 큐레이션 방법론을 제안하는 연구이다. 이러한 관점은 모델 구조를 바꾸기보다 데이터의 질과 구성을 개선하는 데 집중하는 데이터 중심(Data-Centric) 접근과 흐름을 같이한다.

이 관점에서 본 연구의 구성 요소는 다음과 같이 정리된다. 학습 모델인 NVIDIA PilotNet과 검증 환경인 CARLA 시뮬레이터는 본 연구가 개선하거나 발명한 대상이 아니라, 제안 방법론의 효과를 측정하기 위해 채택한 표준 도구이다. 어떤 모델을 사용하더라도 본 연구가 제안하는 큐레이션 방법론은 동일하게 적용될 수 있으며, PilotNet은 구조가 단순하고 널리 검증되어 있어 도구로서 적합하다. 본 연구의 기여는 이 고정된 도구 위에서 어떤 데이터를 학습에 제공할 것인가를 결정하는 전략, 곧 CBHEM에 있다.

2. 제안 시스템 개요: 폐루프 아키텍처

제안 방법론을 실제로 작동시키기 위한 전체 골격이 VITAL의 폐루프(closed-loop) 아키텍처이다. 이 구조의 가장 큰 특징은 자율주행 버스의 운행과 학습이 분리된 일회성 과정이 아니라, 한 주기의 운행이 곧 다음 주기의 학습으로 이어지는 순환 구조라는 점이다. 이는 버스가 동일한 노선을 매일 반복 운행한다는 특성을 적극적으로 활용한 것으로, 같은 구간을 반복 진단하고 보강할 수 있게 한다.

시스템은 운행을 시작하기 전 한 차례만 수행하는 사전 정의 단계와, 매 주기 반복되는 다섯 단계의 사이클로 구성된다(그림 1). 사전 정의 단계에서는 자율주행 버스가 겪는 실패 양상을 세 가지 범주, 즉 급정거·차선변경·곡선차로로 구분한다. 이 세 범주는 실제 시범 운행에서 빈번히 보고되는 문제를 바탕으로 정의된 것으로, 이후 데이터를 분류하고 관리하는 기준이 되며 특정 실패 유형에 학습이 과도하게 치우치는 것을 막는 균형의 단위가 된다. 이 단계는 시스템 구축 시 한 번만 수행되며, 이후의 사이클에서는 반복되지 않는다. 사전 정의가 끝나면 다음의 다섯 단계가 매 주기 반복된다.



< 그림 1 > VITAL 페루프 아키텍처

1) Drive — 운행 및 데이터 수집. 보수적 주행 정책이 가상 노선을 운행하면서 전방 카메라 영상과 차량 주행 신호(속도·조향각·가감속·차간거리 등)를 수집한다. 이때 보수적 정책은 신호등을 전방 카메라로만 인지하므로, 인식 범위·시야각의 제약과 확률적 미검출로 인해 정상적인 흐름에서도 과도한 망설임과 정지가 유발된다. 차량의 속도 변화, 조향 각도, 제동 정도 등이 미리 설정한 임계값을 넘어서면 시스템은 해당 순간을 세 실패 범주 중 하나로 자동 분류하여 기록한다. 이때 사용하는 임계값은 실제 자율주행 버스의 주행 데이터에서 도출하며, 자세한 내용은 IV장에서 다룬다.

2) Diagnose — 취약 구간 진단. 보수적 정책이 노출한 실패는 어느 구간이 취약한지를 알려주는 진단 신호로 활용된다. 당초 본 연구는 안전관리원의 수동 개입(takeover)이 발생한 순간의 조작값을 학습 후보로 직접 활용하는 방안을 계획하였다. 그러나 시뮬레이터 환경에서 연구자가 직접 수행하는 수동 조작은 실차의 숙련된 안전관리원과 달리 거칠고 일관성이 낮아, 동일한 상황에서도 매번 다른 조작값이 기록되는 등 지도 신호로서의 품질을 보장하기 어려웠다. 이에 본 연구는 개입의 역할을 재정의하였다. 즉 개입(실패 신호)은 보강이 필요한 약점 지점을 가리키는 진단 원리로만 계승하고, 그 지점의 정답 조작값은 다음 단계에서 규칙 기반 expert가 일관되게 제공하도록 역할을 분리하였다. 이는 CARLA 기반 모방학습 연구들이 사람의 시연 대신 특권 정보(privileged information)를 가진 규칙 기반 전문가 정책을 정답 출처로 사용하는 확립된 관행과도 부합한다.

3) Curate — expert 정답 확보 및 데이터 교체. 노선을 정류장 단위 구간으로 나누고, 각 구간의 실패율을 통계적으로 비교하여 유의하게 성능이 낮은 취약 구간을 선별한다. 선별된 취약 구간에 대해서는 규칙 기반 expert가 정답 주행 데이터를 제공한다. expert는 차량 간 협력 통신(C-ITS)으로 신호 정보를 신뢰성 있게 수신하고 안전 차간거리를 유지하는 모델로, 보수적 정책이 카메라 인지의 한계로 실패한 바로 그 구간에서 올바른 주행을 일관되게 시연한다. 이렇게 확보한 정답 데이터를 세 범주별 데이터 슬롯에 분배하여 기존 데이터의 일부와 교체하며, 이 교체 과정의 구체적인 방법이 본 연구의 핵심인 CBHEM이다. 자세한 내용은 3절에서 다룬다.

4) Learn — 모델 재학습. 갱신된 데이터셋을 사용하여 주행 제어 모델(PilotNet)을 다시 학습시킨다. 이때 학습은 expert의 올바른 조작을 그대로 모방하도록 하는 행동 복제(Behavior Cloning) 방식으로 이루어진다. 새롭게 보강된 취약 구간의 정답 데이터가 반영되면서, 모델은 이전 주기에서 실패했던 상황을 더 잘 처리할 수 있도록 개선된다.

5) Evaluate — 학습 효과 정량 측정. 재학습된 모델의 성능을 측정하여 학습이 실제로 효과가 있었는지를 정량적으로 확인하고, 그 결과를 다음 주기에서 취약 구간을 다시 선별하는 기준으로 삼는다. 설계상 이 단계는 재학습된 모델의 재주행 점수로 측정하는 것이 이상적이나, 본 연구의 검증에서는 데이터 전략의 기여를 분리하기 위해 IV장에서 기술하는 오픈루프(open-loop) 평가, 즉 별도 검증 주행 데이터에 대한 예측 오차 측정으로 이를 수행한다.

이상의 다섯 단계가 반복되면서, 자율주행 버스는 보수적 정책의 실패가 잦았던 구간을 점차 안정적으로 주행할 수 있게 된다. 이는 곧 안전관리원 동승 단계에서 완전 무인 단계로 나아가는 점진적 진화의 경로를 의미한다.

3. 구간 선택적 데이터 큐레이션(CBHEM)

앞 절의 큐레이션 단계는 단순히 새 데이터를 추가하는 작업이 아니다. 본 연구는 이 과정을 CBHEM(Class-Balanced Hard Example Mining)이라는 데이터 큐레이션 방법으로 정의하며, 이것이 본 연구의 가장 핵심적인 기여이다. CBHEM은 어려운 예시를 집중적으로 학습하는 OHEM과 범주 간 균형을 유지하는 Class-Balanced Learning을 결합한 기법으로, ① 어느 구간을 보강할 것인가(구간 선별), ② 그 구간에서 어떤 범주를 갱신할 것인가(범주 선별), ③ 슬롯 내부에서 무엇을 어떻게 교체할 것인가(예시 선별)의 세 단계로 작동한다.

단순 보강이 아닌 큐레이션이 필요한 이유는 다음과 같다. 자율주행 모델을 개선하는 가장 단순한 방법은 실패한 데이터를 많이 모아 그 비중을 높여 학습시키는 것이다. 그러나 이러한 단순 가중 학습은 특정 실패 유형의 데이터가 학습을 지배하게 되어, 그 상황에는 강해지는 대신 이전에 잘 처리하던 다른 상황을 잊는 파국적 망각(catastrophic forgetting)을 야기한다. CBHEM은 범주 간 비율을 고정한 채 각 범주 내부의 내용만 선별하여 교체함으로써, 약점 보강과 망각 방지를 동시에 달성한다.

1) 구간 선별 — 통계적 약점 식별

취약 구간의 선별은 임의의 하위 비율을 자르는 방식이 아니라 통계적 유의성에 근거한다. 노선을 정류장 단위로 나눈 각 구간 s에 대해 실패율을 산출한 뒤, 전체 구간 실패율의 평균 μ와 표준편차 σ로부터 임계값을 정의한다.

$$\theta = \mu + \delta\sigma \tag{1}$$

이어서 각 구간의 실패율에 대해 부트스트랩(bootstrap) 재표집을 수행하여 90% 신뢰구간의 하한을 추정하고, 이 하한이 임계값을 초과하는 구간만을 통계적으로 유의한 취약 구간 집합 W로 선정한다.

$$W = \{\, s \mid r_s^{lo} > \theta \,\} \tag{2}$$

여기서 단순 평균이 아니라 신뢰구간의 하한을 사용하는 것이 핵심이다. 어떤 구간의 표본 수가 적으면 실패율이 우연히 높게 관측될 수 있는데, 이 경우 점 추정값은 크더

라도 신뢰구간이 넓어 하한은 낮게 나온다. 하한을 기준으로 삼으면 이러한 표본 부족에 의한 거짓 약점을 자동으로 배제할 수 있어, 반복 관측해도 일관되게 성능이 낮은 구간만 선별된다. 한편 주행 중 발생하는 충돌 사건은 시뮬레이터 환경의 특성에서 기인할 가능성이 있어 약점 판정에 반영하지 않으며, 이에 관한 논의는 VI장에서 다룬다. 실험에서는 δ=1.0을 사용하였다.

2) 범주 선별 — 기준 대비 과락 판정

약점 구간이 정해지면, 그 구간 안에서 세 실패 범주 중 어떤 범주를 이번 주기에 갱신할지를 결정한다. 이때 단순히 실패 건수가 많은 범주를 고르면 표본이 풍부한 범주가 매번 선택되어 균형이 무너진다. 예를 들어 차선변경은 곡선차로보다 절대 실패 수가 많을 수 있으나, 이는 그 상황이 더 자주 발생하기 때문일 뿐 반드시 더 심각한 약점이라고 볼 수 없다. 따라서 본 연구는 각 범주의 실패율을, 그 범주가 정상 조건에서 자연히 발생하는 기준 허용치 τ로 정규화한 상대 위험도(Category Risk Factor, CRF)를 판정 기준으로 삼는다.

$$CRF_c = r_c / \tau_c \tag{3}$$

기준 허용치 τ는 교통·신호가 없는 정상 운행 조건의 검증 주행에서 각 범주가 자연스럽게 발생하는 비율로 추정한다. CRF가 임계 k 이상인 범주를 정상 수준을 유의하게 초과하여 실패하는 과락 범주 F로 판정하되, 한 주기에 갱신하는 범주 수는 상한 cap으로 제한하여 점진적이고 안정적인 갱신을 보장한다.

$$F = \{\, c \mid CRF_c \geq k \,\} \ \text{(상위 cap개)} \tag{4}$$

이로써 표본의 절대량이 아니라 기준 대비 초과 정도를 척도로 삼게 되어, 표본이 적은 소수 범주라도 비정상적으로 자주 실패하면 우선 보강 대상이 된다. 실험에서는 k=2.0, cap=2를 사용하였다.

3) 예시 선별 — 균형 고정과 어려운 예제 교체

세 번째 단계에서 CBHEM은 세 실패 범주에 대응하는 세 개의 데이터 슬롯을 두고, 각 슬롯이 전체의 1/3씩을 차지하도록 비율을 항상 고정한다. 전체 학습셋 크기를 3N이라 하면 각 범주 슬롯은 N개로 고정되며, 새로운 데이터가 유입되어도 이 비율은 변하지 않는다. 이 고정이 파국적 망각을 원천적으로 차단하는 장치이다. 특정 범주의 데이터가 아무리 많아도 슬롯 크기를 넘겨 학습을 지배할 수 없기 때문이다. 과락 범주의 슬롯을 채울 때에는 다음 세 가지 원리를 적용한다.

(1) 유효 표본 수 기반 배분. 과락 범주 안에서도 구간마다 표본 수가 다르므로, 표본이 많은 구간이 슬롯을 독점하지 않도록 배분을 조정한다. 본 연구는 Cui 등[7]이 제안한 유효 표본 수(Effective Number of Samples) 개념을 도입한다. 표본이 n개인 구간의 유효 표본 수는 다음과 같이 정의된다.

$$E_n = (1 - \beta^n) / (1 - \beta) \tag{5}$$

$$w_s = 1 / E_{n\_s} = (1 - \beta) / (1 - \beta^{n\_s}) \quad (5\text{-}1)$$

β는 0과 1 사이의 상수로, β가 1에 가까울수록 유효 표본 수는 단순 표본 수 n에 가까워지고 β가 작을수록 표본 증가에 따른 기여가 빠르게 포화한다. 각 구간의 배분 가중치를 유효 표본 수의 역수 w_s로 두면, 표본이 많은 구간일수록 가중치가 작아져 배분량이 줄어든다. 이는 Cui 등[7]의 Class-Balanced 가중치 (1−β)/(1−β^n)와 동일한 형태로, 동일한 장면이 중복 수집되어도 그 정보량이 선형으로 늘지 않는다는 관찰에 근거하여 다수 구간으로의 쏠림을 억제한다. 실험에서는 β=0.999를 사용하였다.

(2) 약점 구간 최소 할당 보장. 정작 보강이 필요한 약점 구간의 데이터가 배분 과정에서 밀려나면 큐레이션의 목적이 무색해진다. 이를 막기 위해, 약점 구간 s∈W에 대해서는 그 구간의 자연 비율에 보정 계수 ρ를 곱한 양을 최소 할당량으로 보장한다. 즉 구간 s의 배분량은 유효 표본 수 기반 배분값과 이 최소 할당량 중 큰 값으로 결정된다.

$$q_s = \max(\, q_s^{E} \,,\, \mathrm{round}(\mathrm{natural}_s \times \rho) \,) \quad (6)$$

여기서 q 위첨자 E는 식 (5-1)의 가중치에 따른 배분량, natural은 구간 s의 자연 비율이다. 보정 계수 ρ>1은 약점 구간을 자연 비율보다 상향 보강하겠다는 의도를 수치로 표현한 것으로, 실험에서는 ρ=1.5를 사용하였다.

(3) 어려운 예제와 무작위 표본의 혼합. 각 구간에 배정된 만큼의 프레임을 고를 때, 학습 난이도를 나타내는 손실값(loss)이 높은 예시를 우선한다. 손실값이 낮은 예시는 모델이 이미 잘 처리하고 있어 더 학습할 가치가 적은 반면, 손실값이 높은 예시는 모델이 어려워하는 지점이므로 학습 효율이 높다. 다만 어려운 예시만으로 슬롯을 채우면 특정 상황에 과적합되고 데이터 다양성이 훼손될 수 있다. 이에 배정량 k개 중 비율 h만큼만 손실값 상위에서 선택하고 나머지는 무작위로 채우는 혼합 방식을 사용한다.

$$n_{\mathrm{hard}} = \mathrm{round}(k \cdot h) \quad (7)$$

즉 상위 n_hard개는 어려운 예제로, 나머지는 무작위 표본으로 구성한다. h=1이면 순수 하드 마이닝(OHEM과 동일)이 되고, h가 낮을수록 다양성이 강조된다. 실험에서는 h=0.5로 균형을 두었다. 비과락 범주의 슬롯은 갱신 대상이 아니므로, 해당 범주의 데이터를 균형 크기 N에 맞추어 무작위로 유지한다. 이상의 세 원리를 하나의 절차로 정리하면 알고리즘 1과 같으며, 전체 파이프라인을 그림 2에 나타내었다.

[알고리즘 1] CBHEM 데이터 큐레이션

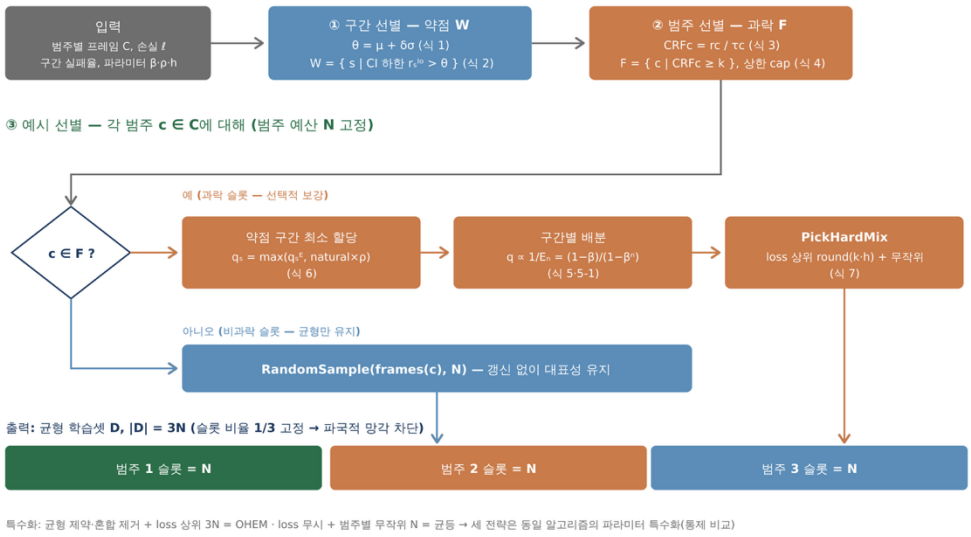
```
입력: 범주별 프레임 C = {급정거, 차선변경, 곡선차로}, 약점 구간 W,
      과락 범주 F, 범주 예산 N, 손실함수 loss, 파라미터 β, ρ, h
출력: 균형 학습셋 D  (|D| = 3N)

D <- 공집합
for 각 범주 c in C do
    if c in F then                          // 과락 범주: 선택적 보강
        slot <- 공집합
        for 각 구간 s in c do
            q_s <- E_s(식 5)에 비례한 배분량
```

```
            if s in W then                  // 약점 구간 최소 할당 (식 6)
                q_s <- max(q_s, round(natural_s * ρ))
            slot <- slot + PickHardMix(frames(c,s), q_s, h)
        if |slot| < N then slot <- slot + RandomFill(c, N-|slot|)
    else                                    // 비과락 범주: 균형만 유지
        slot <- RandomSample(frames(c), N)
    D <- D + slot
return D


함수 PickHardMix(F, k, h):                  // 어려운 예제 + 무작위 혼합 (식 7)
    n_hard <- round(k * h)
    hard   <- loss 기준 상위 n_hard개
    rand   <- (F - hard)에서 무작위 (k - n_hard)개
    return hard + rand
```

알고리즘 1은 본 연구의 비교군을 하나의 틀로 통합한다. 균형 제약(1/3 고정)과 최소 할당·혼합을 모두 제거하고 전체에서 손실값 상위 3N개만 취하면 균형 없는 하드 마이닝, 즉 OHEM 비교군이 된다. 반대로 손실값을 무시하고 각 범주에서 무작위로 N개씩 취하면 균등 학습 비교군이 된다. 이처럼 CBHEM·OHEM·균등 학습이 동일한 알고리즘의 파라미터 특수화로 표현되므로, 세 전략의 비교는 데이터 구성 전략의 차이만을 분리하여 측정하는 통제 비교가 된다.



<그림 2> CBHEM 데이터 큐레이션 파이프라인(알고리즘 1)

이러한 범주 구분은 데이터 큐레이션 단계에서만 사용되며, 주행 제어 모델은 범주를 알지 못한 채 영상·센서와 조작값의 쌍만 학습한다. 따라서 모델을 교체하더라도 CBHEM 방법론은 그대로 재사용할 수 있다.

4. 학습 모델 구현

주행 제어 모델은 NVIDIA PilotNet[1]을 기반으로 구현한다. PilotNet은 200×66 크기의 RGB 영상을 입력받아 조작값을 회귀하는 End-to-End 합성곱 신경망으로, 구조가 단순하고 널리 검증되어 있어 데이터 전략의 효과를 분리 측정하는 도구로 적합하다.

본 연구는 원 구조의 카메라 단독 입력을 카메라 영상과 차량 상태 센서(속도·종가속도·조향)의 결합 입력으로 확장한다. 보수적 정책의 대표적 실패인 불필요한 급제동은

장면 정보만으로는 판별이 어렵고, 눈앞의 장면(카메라)과 차량의 거동(센서)을 함께 관측하여 비어 있는 길에서의 급제동과 같은 장면-거동 간 모순을 포착할 수 있어야 억제된다. 이러한 카메라·상태 결합 입력은 CILRS[8] 등 End-to-End 자율주행 연구에서 확립된 표준 구성이다. 구현상으로는 CNN이 추출한 영상 특징 벡터에 센서 벡터를 연결(concatenation)하여 완전연결 계층에 전달하며, 모델은 조향·가속·제동의 세 조작값을 출력한다. 학습은 expert의 조작값을 지도 신호로 하는 행동 복제 방식으로 수행하고, 데이터 범주 정보는 학습 과정에 일절 노출하지 않는다.

Ⅳ. 실험 설계

본 장에서는 제안 방법의 효과를 통제된 조건에서 측정하기 위한 실험 설계를 기술한다. 1절에서 시뮬레이션 환경과 노선을, 2절에서 데이터 수집과 실패 범주 분류를, 3절에서 평가 지표와 비교군을, 4절에서 통계적 유의성 검정 방법을 제시한다.

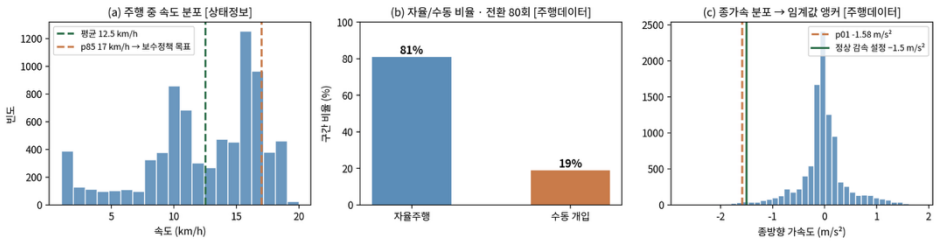
1. 실험 환경 및 노선 구성

실험 환경은 CARLA 0.9.15 시뮬레이터[8]의 Town10HD 맵에 구축하였다. 맵의 도로 구조를 활용하여 정류장 8개를 배치한 순환 노선을 구성하고, 정류장을 기준으로 노선을 8개 구간으로 분할하였다. 이 노선은 특정 실노선의 지리를 재현한 것이 아니라, 정류장 정차와 곡선·직선 주행이 반복되는 버스 운행의 전형적 구조를 갖추도록 설계한 것이며, 분할된 각 구간은 Ⅲ장에서 기술한 구간 선별의 단위가 된다.

실험에는 서로 다른 인지 능력을 가진 세 주체가 등장한다. **보수적 주행 정책**(진단)은 신호등을 전방 카메라로만 인지하여 확률적 미검출이 발생하며, 판타G 운행 데이터의 순항 속도(p85)에 맞추어 목표 순항 17km/h로 주행하는 진단용 주체이다. **expert**(정답)는 차량 간 협력 통신(C-ITS)으로 신호 정보를 신뢰성 있게 수신하고 안전 차간거리를 유지하며, 목표 순항 30km/h로 주행하는 정답 제공 주체이다. **PilotNet**(학습 대상)은 이 두 주체 사이에서 expert의 조작을 모방하여 학습된다. 보수적 정책과 expert의 인지 비대칭이 곧 실패와 정답의 구도를 형성하며, 두 목표 속도의 격차가 학습으로 좁힐 개선 여지에 해당한다.

실차의 특성은 경기도 판교 판타G버스의 2024년 공개 주행 데이터[11]로 두 가지 방식에 반영하였다(그림 3). 첫째, 실차의 보수적 거동을 보수적 정책의 설계 근거로 삼았다. 공개 데이터에서 버스는 주행 중 평균 약 12.5km/h의 저속으로 운행하며, 약 19%의 구간이 수동 개입 상태였고 자율에서 수동으로의 제어권 전환이 80회 관측되었다. 이는 개입이 실차 운행에서 실재하며 체계적으로 발생함을 보여주는 근거로, 본 연구가 개입을 약점 진단의 신호로 삼는 출발점이 된다. 둘째, 주행 신호의 분포를 실패 범주 임계값과 거동 파라미터의 현실 앵커로 사용하였다. 정상 주행의 감속은 종방향 가속도의 하위 백분위(약 -1.58 m/s²)에 맞추어 1.5 m/s²로, 비상 감속은 3.0 m/s², 최대 저크는 1.5 m/s³로 설정하였고, 곡선 주행의 횡방향 가속도 한계는 실데이터의 p95(약 0.62 m/s²)를 참조하였다. 이 설정값들은 적응순항제어(ACC) 성능을 규정한 국제 표준 ISO 1

5622:2018[10]이 제시하는 상한(감속 3.5 m/s², 저크 2.5 m/s³)보다 보수적인 범위에 있다. 다만 1Hz 주기의 텔레메트리 로그는 급격한 거동을 평활하고 카메라 영상을 포함하지 않으므로, 실데이터는 거동·임계의 근거로만 활용하고 학습 데이터로는 사용하지 않았다. 주변 교통은 차량 24대와 저속 선행차 시나리오로 구성하고 신호 주기를 고정하였으며, 보수적 정책 수집과 expert 수집에 동일한 교통 설정을 적용하였다.



<그림 3> 판타G버스 공개 주행 데이터 분석

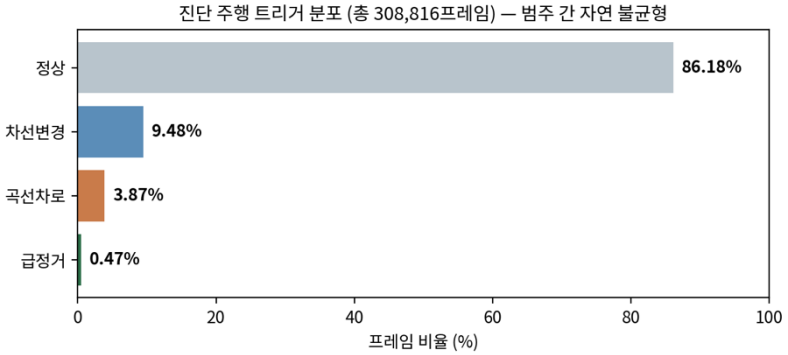
2. 데이터 수집 및 실패 범주 분류

취약 구간의 진단을 위해 보수적 정책으로 서로 다른 난수 시드의 주행 20회를 수행하여 총 308,816프레임을 수집하였다. 각 프레임은 주행 신호를 기준으로 세 실패 범주 중 하나로 자동 분류되며, 분류 기준은 표 1과 같다. 한 프레임이 여러 조건을 충족하면 차선변경, 곡선차로, 급정거의 우선순위에 따라 배타적으로 단일 범주에 배정된다.

<표 1> 실패 범주 자동 분류 기준

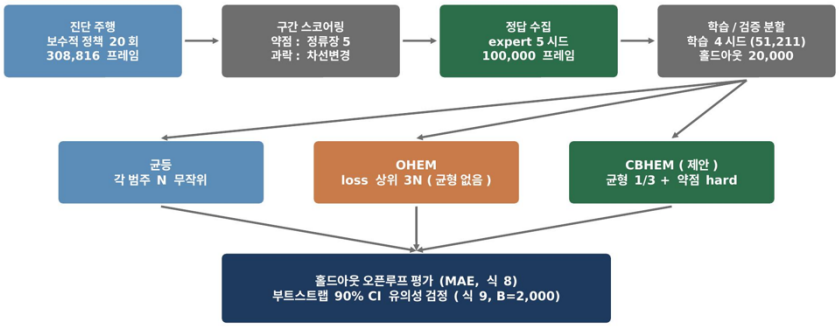
범주	판정 조건
차선변경	전방 20m 이내에 15km/h 미만의 선행차를 추종하며 차로 변경을 망설이는 상태
곡선차로	곡선 구간을 주행하는 상태
급정거	제동 입력이 우세하거나(brake>0.2) 정지에 접근하는 상태

이러한 자동 분류를 활용하면 사람이 데이터를 일일이 검토하지 않고도 어느 구간에서 어떤 문제가 발생했는지 체계적으로 누적할 수 있다. 진단 주행의 트리거 분포는 정상 86.2%, 차선변경 9.5%, 곡선차로 3.9%, 급정거 0.5%로, 실패 범주 간에 큰 자연 불균형이 존재한다. 이 불균형은 특정 범주에 학습이 치우치기 쉬움을 보여주며, 범주 균형을 고정하는 CBHEM의 필요성을 뒷받침한다(그림 4).



<그림 4> 진단 주행의 실패 범주 트리거 분포

구간 스코어링이 취약 구간과 과락 범주로 지목한 지점의 정답을 확보하기 위해, 실패가 집중된 다섯 개의 시드에 대해 expert 주행을 각각 수행하여 시드당 20,000프레임씩 총 100,000프레임의 정답 데이터를 수집하였다. expert 주행은 보수적 정책과 동일한 노선·교통 설정에서 시드만 달리하여, 보수적 정책이 실패한 바로 그 상황에 대한 정답을 대응시켰다. 이 가운데 네 시드를 학습에 사용하고 한 시드를 검증용 홀드아웃으로 분리하였다. 이상의 전체 실험 절차를 그림 5에 나타내었다.



< 그림 5 > 실험 파이프라인

세 데이터 전략의 학습셋은 동일한 크기로 구성한다. 학습에 사용한 네 시드의 정답 데이터를 범주별로 분류하면 급정거 9,245, 차선변경 13,380, 곡선차로 28,586, 합계 51,211프레임이 된다. 세 범주 중 표본이 가장 적은 급정거를 기준으로 범주당 예산을 N=9,245로 정하고, 전체 학습셋 크기를 3N=27,735프레임으로 고정하였다. 균등·OHEM·CBHEM 세 전략이 모두 같은 크기의 학습셋을 사용하므로, 성능 차이는 오직 데이터 구성 전략의 차이에서만 비롯된다.

3. 평가 지표 및 비교군

데이터를 학습셋과 검증셋으로 나눌 때에는 프레임을 무작위로 분할하지 않고 주행 에피소드 단위로 분리하였다. 연속한 프레임은 서로 매우 유사하므로 무작위로 섞으면 학습셋과 검증셋에 거의 같은 장면이 포함되어 성능이 부풀려지기 때문이다. 이에 네 시드의 주행을 학습에 사용하고, 학습에 쓰이지 않은 별도의 시드 주행 20,000프레임을 검증용 홀드아웃으로 남겼다. 이 홀드아웃에는 약점 구간의 실패 상황이 포함되어 있어 약점 개선을 직접 측정할 수 있다.

제안 방법의 효과는 학습 전(보수적 정책)을 개선 폭의 기준선으로 두고, 균등·OHEM·CBHEM 세 전략을 동일한 도구 위에서 비교하여 확인한다.

평가 지표는 홀드아웃의 각 프레임에서 모델이 출력한 조향·가속·제동 세 조작값과 expert 정답 사이의 평균절대오차(MAE)이다. 이는 모델을 실제로 주행시키지 않고 검증 데이터에 대한 예측 오차를 재는 오프루프 평가에 해당한다. 프레임 i의 예측을 $(\hat{s}, \hat{t}, \hat{b})$, 정답을 (s, t, b)라 하면 오차와 MAE는 다음과 같이 정의된다.

$$e_i = (1/3)(|\hat{s}_i - s_i| + |\hat{t}_i - t_i| + |\hat{b}_i - b_i|) \quad (8\text{-}1)$$

$$MAE(D) = (1/|D|)\ \Sigma_i\ e_i \quad (8)$$

측정은 전체 홀드아웃뿐 아니라 약점 구간(정류장5의 차선변경), 망각 점검을 위한 개별 범주(급정거·곡선차로), 조작 채널(조향·가속·제동), 그리고 가감속 성향의 expert 근접도로 나누어 수행하여, 약점 개선과 기존 능력 유지를 함께 확인한다. 아울러 각 전략이 자신의 모델 손실값으로 어려운 예제를 다시 선별하여 재학습하는 과정을 반복하여, 큐레이션을 반복했을 때의 성능 안정성을 관찰한다. 이 반복은 고정된 데이터 풀 안에서 이루어지므로, 새로운 실패가 계속 유입되는 완전한 폐루프와는 구분된다.

4. 통계적 유의성 검정

점 추정된 MAE 값만으로는 두 전략의 차이가 우연에 의한 것인지 판단할 수 없다. 이를 검정하기 위해 부트스트랩(bootstrap) 신뢰구간을 도입한다. 홀드아웃 프레임을 복원추출로 B회 재표집하여 각 재표본의 MAE 분포를 구하고 그 90% 신뢰구간을 산출한다. 두 전략 A와 B를 비교할 때에는 동일한 재표본을 양쪽에 적용하는 대응 방식으로 차이의 분포를 구하고, 그 신뢰구간이 0을 포함하지 않으면 통계적으로 유의한 차이로 판정한다.

$$\Delta_b = \mathrm{MAE}_A(\boldsymbol{\mathcal{B}}_b) - \mathrm{MAE}_B(\boldsymbol{\mathcal{B}}_b), \quad b = 1, \ldots, B \tag{9-1}$$

$$\text{유의} \Leftrightarrow 0 \notin [\, Q_{0.05}(\Delta),\ Q_{0.95}(\Delta) \,] \tag{9}$$

실험에 사용한 방법론·검정 파라미터를 표 2에 정리하였다. 이 신뢰구간은 평가 표본의 불확실성을 반영한 것으로, 학습 시드에 따른 변동까지 포괄하는 다중 시드 반복 검증은 향후 연구에서 견고성을 강화하는 방향으로 확장할 수 있다.

< 표 2 > 방법론 및 검정 파라미터

파라미터	값	의미
δ	1.0	약점 임계 μ+δσ
k, cap	2.0, 2	과락 CRF 임계·사이클 상한
τ (급정거/차선변경/곡선차로)	0.02 / 0.05 / 0.05	범주별 기준 허용치
β	0.999	유효 표본 수
ρ	1.5	약점 구간 최소 할당 배수
h	0.5	어려운 예제 혼합 비율
N, 3N	9,245 / 27,735	범주 예산·학습셋 크기
B, 신뢰수준	2,000 / 90%	부트스트랩 반복·신뢰수준

Ⅴ. 실험 결과 및 분석

본 장에서는 Ⅳ장의 실험 설계에 따라 수행한 결과를 제시한다. 1절에서 약점 구간과 과락 범주를 식별하고, 2절에서 데이터 전략별 통제 비교로 전체 성능을 확인한다. 3절

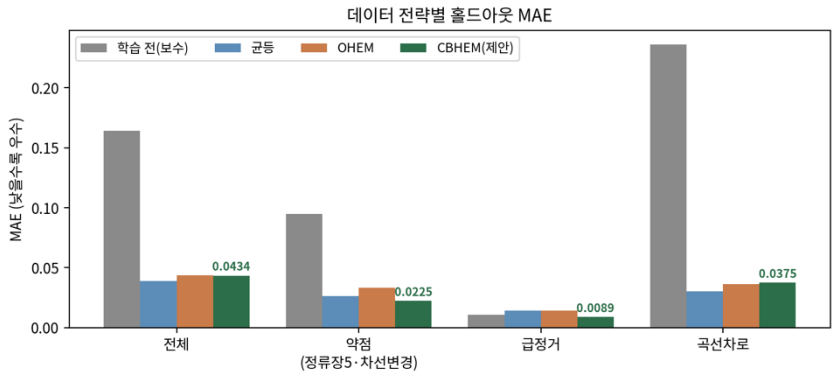
에서 본 연구의 핵심인 약점 개선과 파국적 망각 방지를 함께 분석하고, 4절에서 반복 사이클의 안정성을, 5절에서 통계적 유의성을 종합한다. 모든 오차는 홀드아웃 주행(학습 미사용)에서 측정한 평균절대오차(MAE)이며, 낮을수록 expert 정답에 가깝다.

1. 약점 구간·과락 범주 식별

보수적 정책의 진단 주행을 구간별로 스코어링한 결과, 여덟 개 정류장 구간 중 정류장5가 통계적으로 유의한 약점 구간으로 선별되었다. 이 구간은 실패율의 90% 신뢰구간 하한이 임계값 μ+δσ를 넘어, 표본 부족에 의한 우연이 아니라 반복적으로 성능이 낮은 구간임이 확인되었다. 이어 정류장5 내부의 범주별 상대 위험도(CRF)를 산출한 결과, 차선변경이 CRF 3.34로 임계 k=2.0을 크게 초과하여 과락 범주로 판정되었다. 즉 정류장5의 차선변경 상황이 이번 큐레이션의 보강 대상이며, 이후의 모든 비교는 이 약점을 중심으로 이루어진다.

2. 데이터 전략별 통제 비교

동일한 PilotNet 모델과 동일한 크기(3N=27,735)의 학습셋 위에서 데이터 전략만 균등·OHEM·CBHEM으로 달리하여 홀드아웃 전체 MAE를 비교하였다(그림 6). 세 전략 모두 학습 전 보수적 정책(전체 0.1641)에 비해 큰 폭으로 오차를 낮추어, 취약 구간의 정답을 학습하는 접근 자체가 효과적임을 보였다.



<그림 6> 데이터 전략별 홀드아웃 MAE

전체 MAE에서는 균등(0.0389)이 CBHEM(0.0434)보다 근소하게 낮다. 이는 홀드아웃에서 곡선차로 표본이 다수(n=6,894)를 차지하는데, 균등 전략이 이 다수 범주에 자연히 더 많이 학습하여 곡선차로 오차(0.0304)를 낮춘 데서 비롯된 결과이다. 그러나 이 전체 평균의 우위는 소수 범주와 약점 구간을 희생한 대가이며, 본 연구가 주목하는 지점은 바로 그 희생된 영역이다. 다음 절에서 이를 분석한다.

3. 약점 구간 개선 및 파국적 망각 방지

본 연구의 핵심 주장은 CBHEM이 약점 구간을 개선하면서도 다른 범주를 잊지 않는다는 것이다. 두 축을 함께 확인한다.

첫째, 약점 구간의 개선이다. 정류장5·차선변경의 MAE는 학습 전 0.0950에서 CBHEM

0.0225로 낮아져 76.3%의 개선을 보였으며, 큐레이션이 없는 균등(0.0263)과 비교해도 CBHEM이 더 낮았다. 어려운 예제만 강조하는 OHEM(0.0335)은 오히려 균등보다 나빠, 약점 보강에는 균형 유지가 결정적임을 보였다.

둘째, 파국적 망각의 방지이다. 소수 범주인 급정거에서 CBHEM(0.0089)은 균등(0.0142)과 OHEM(0.0141)을 모두 앞섰다. 균등은 자연 분포를 따르므로 급정거를 상대적으로 적게 학습하고, OHEM은 손실값이 큰 곡선차로에 쏠려 급정거를 사실상 버린다. 반면 CBHEM은 범주 비율을 1/3로 고정하여 급정거 슬롯을 온전히 확보하므로, 소수 범주를 잊지 않는다. 이것이 균형 큐레이션의 직접적 효과이다.

이 개선은 조작 채널과 거동에서도 확인된다(표 3). 약점 구간에서 CBHEM은 가속(throttle) 오차를 균등보다 낮추었고(0.0579 대 0.0662), 무엇보다 가감속 성향(throttle−brake)이 expert 정답에 가장 근접하였다. expert의 성향(+0.095)을 기준으로 CBHEM은 +0.101로 그 차이가 0.006에 불과하여, 세 전략 중 expert의 거동을 가장 정확히 재현하였다. 이는 균등(차이 0.019)·OHEM(차이 0.066)보다 뚜렷이 우수하며, 학습 전 보수 정책(차이 0.234)과는 큰 격차를 보인다. 즉 CBHEM은 약점 구간에서 오차를 낮출 뿐 아니라 조작의 거동 자체를 정답에 가장 가깝게 학습하였다.

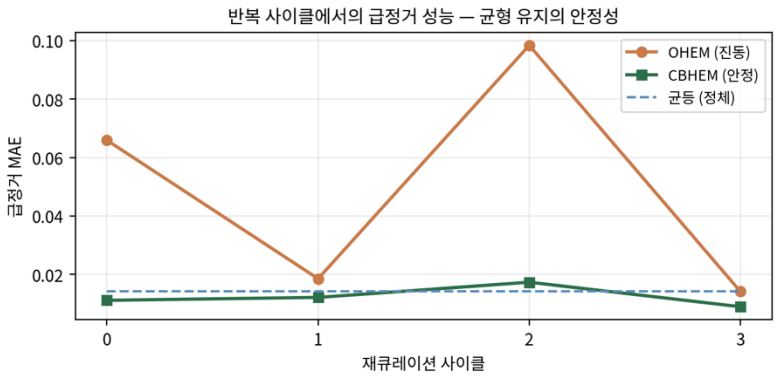
< 표 3 > 약점 구간 거동 분석 (정류장5·차선변경)

지표	학습 전	균등	OHEM	CBHEM
가속(throttle) MAE	0.2750	0.0662	0.0881	0.0579
제동(brake) MAE	0.0100	0.0098	0.0082	0.0074
가감속 성향 (expert=+0.095)	+0.329	+0.115	+0.161	+0.101
expert와의 차이	0.234	0.019	0.066	0.006

이로써 CBHEM은 약점 구간을 학습 전 대비 크게 개선하면서, 동시에 소수 범주를 보존하고 거동을 정답에 근접시켰다.

4. 반복 사이클 안정성

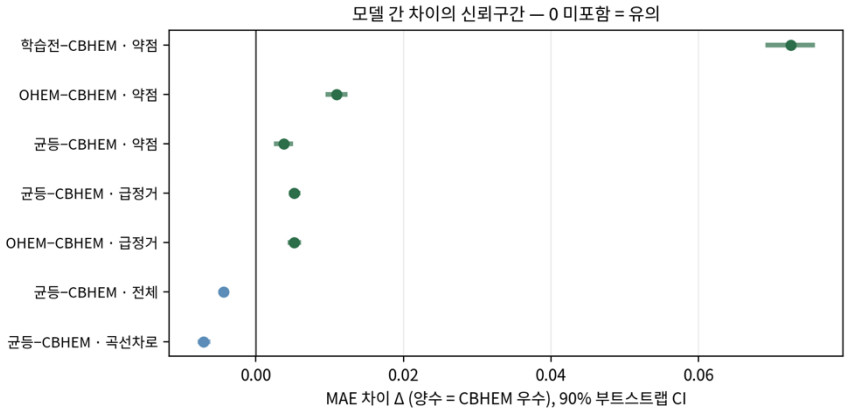
큐레이션을 한 번만 수행하는 것을 넘어, 각 전략이 자신의 손실값으로 어려운 예제를 다시 선별·재학습하는 과정을 반복했을 때의 거동을 관찰하였다(그림 7). 급정거 범주에서 그 차이가 뚜렷하게 나타났다. OHEM은 사이클에 따라 0.066, 0.019, 0.098, 0.014로 크게 진동하였는데, 이는 손실값 상위만 반복 선별하는 과정에서 학습 대상이 사이클마다 요동치기 때문이다. 반면 CBHEM은 0.011, 0.012, 0.017, 0.009로 0.009에서 0.017의 좁은 범위에서 낮게 유지되었고, 균등은 큐레이션이 없으므로 0.0142에 정체하였다. 즉 CBHEM은 반복 큐레이션에서도 소수 범주 성능을 안정적으로 유지한다. 다만 이 반복은 고정된 데이터 풀 안에서 이루어진 것으로, 새로운 실패가 계속 유입되는 완전한 폐루프에서의 장기 거동은 향후 연구로 남는다.



<그림 7> 반복 재큐레이션 사이클의 급정거 MAE 궤적

5. 통계적 유의성 종합

점 추정된 차이가 우연이 아님을 확인하기 위해, 각 비교를 90% 부트스트랩 신뢰구간(n_boot=2,000, 대응 방식)으로 검정하였다(그림 8). 신뢰구간이 0을 포함하지 않으면 유의한 차이이다. 약점 구간에서 CBHEM은 학습 전 대비 +0.0725, 균등 대비 +0.0038, OHEM 대비 +0.0109의 차이를 보였고, 급정거에서는 균등·OHEM 대비 각각 +0.0052로, 모든 구간에서 신뢰구간이 0을 포함하지 않았다.



<그림 8> 모델 간 MAE 차이의 90% 부트스트랩 신뢰구간

약점 구간과 급정거에서 CBHEM의 우위는 모두 통계적으로 유의하였고, 전체·곡선차로에서 균등이 앞서는 것 또한 다수 범주 편중의 결과이다. 종합하면 CBHEM은 정확도의 최댓값 대신 약점 개선·소수 범주 보존·거동의 정답 근접·반복 안정성을 유의하게 달성하였으며, 이는 균형을 유지한 채 약점만 선별 보강한다는 설계가 유효함을 뒷받침한다.

VI. 논의 및 향후 연구

본 장에서는 본 연구가 마주한 시뮬레이션 환경상의 제약을 정리하고, 이를 바탕으로 향후 연구가 나아갈 방향을 제시한다.

1. 시뮬레이션 환경의 한계와 통제

본 연구는 CARLA 시뮬레이터를 검증 환경으로 채택하였다. 시뮬레이터는 물리 엔진의 근사, 센서 노이즈의 재현성 부족, 주변 차량의 규칙 기반 행동 모형 등에서 실도로 환경과 일정한 격차를 지닌다. 이러한 격차는 데이터 큐레이션 방법이 실도로에서 동일한 폭의 개선을 낼지에 대해 완전한 보장을 제공하기 어렵게 만든다. 본 연구는 이 제약을 인지하고 그 영향이 결과에 반영되지 않도록 통제하는 데 주의를 기울였다.

첫째, 시뮬레이터가 유발할 수 있는 인공적 사건은 약점 판정에서 배제하였다. Ⅲ장에서 기술한 바와 같이 주행 중 발생한 충돌 사건은 CARLA의 물리·상호작용 특성에서 기인할 가능성이 있어 약점 구간 선별에 반영하지 않았다.

둘째, 실차의 특성을 방법 설계에 반영하되 데이터의 한계를 명확히 하였다. 판타G버스 데이터는 카메라 영상이 없고 급격한 거동이 평활된 1Hz 텔레메트리이므로, 거동 앵커와 실패 임계값의 근거로만 사용하고 학습 데이터로는 활용하지 않았다.

셋째, 세 전략을 동일한 환경·모델·학습셋 크기 위에서 비교하는 통제 구조를 채택하였다. 시뮬레이터의 근사가 세 전략에 동등하게 작용하므로, 상대 비교의 결론은 시뮬레이션 환경 자체에 대한 의존성이 낮다.

2. 연구의 확장 방향

본 연구가 확립한 균형 큐레이션의 원리는 다음과 같은 방향으로 확장될 수 있다.

첫째, 평가와 검증의 확장이다. 본 연구의 오픈루프 평가를 넘어, 학습된 모델이 차량을 직접 제어하는 폐루프 조건에서 개입 빈도·구간 통과율·승차감을 측정하고, 다중 시드 반복 학습으로 학습 과정의 변동까지 포괄하는 견고성 검증으로 확장할 수 있다.

둘째, 데이터와 적용 범위의 확장이다. 실차 배치 단계에서 축적되는 숙련 안전관리원의 개입 조작을 정답 후보로 재검토하여 시뮬레이션과 실차를 잇고, 단일 약점 사례에 집중한 본 실험을 여러 노선·약점 구간으로 반복 검증하면 제안 방법이 일반적 큐레이션 원리로 자리매김할 수 있다.

이러한 확장은 본 연구가 데이터 구성 층위에서 확립한 균형 큐레이션의 원리를 실제 자율주행 버스의 운용 환경으로 이전하는 자연스러운 경로이며, 안전관리원 동승 단계에서 완전 무인 단계로 나아가는 점진적 진화의 로드맵을 구성한다.

Ⅶ. 결론

1. 연구 요약

본 연구는 자율주행 버스의 보수적 주행 정책이 반복적으로 실패하는 취약 구간을 선별하여 보강하는 데이터 큐레이션 방법론 CBHEM(Class-Balanced Hard Example Mining)을 제안하였다. 개입이 가리키는 실패는 약점 구간을 식별하는 진단 신호로만 활용하고, 정답 조작값은 일관성이 보장된 규칙 기반 expert로부터 확보하는 역할 분리를 통해

, 개입 조작의 품질에 좌우되던 기존 접근의 한계를 해소하였다. CBHEM은 통계적 약점 식별, 기준 대비 과락 판정, 균형 고정과 어려운 예제 교체의 세 단계로 정의되며, OHEM과 균등 학습이 동일 알고리즘의 특수화로 표현되는 통제 비교 구조를 갖는다.

CARLA 시뮬레이터와 PilotNet을 표준 도구로 고정한 통제 비교에서, CBHEM은 약점 구간(정류장5·차선변경)의 오차를 학습 전 대비 76.3% 개선하였고 이는 부트스트랩 신뢰구간 검정에서 통계적으로 유의하였다. 균형을 무시한 OHEM이 소수 범주를 망각한 것과 달리 CBHEM은 급정거 성능을 유의하게 보존하였으며, 가감속 성향은 expert 정답에 가장 근접하였고, 반복 재큐레이션에서도 안정적인 성능을 유지하였다. 이 결과는 전체 학습 풀의 약 54%(27,735/51,211프레임)만으로 달성된 것으로, 데이터의 균형을 유지한 채 약점만 선별 보강하는 전략이 정확도·균형·효율을 함께 달성할 수 있음을 보였다.

2. 기대 효과 및 활용 방안

본 연구의 기대 효과는 세 가지로 정리된다. 첫째, 데이터 효율이다. 실패가 집중되는 구간의 데이터만 선별 보강하므로, 무차별적 데이터 수집·라벨링에 드는 비용을 줄이면서 필요한 성능 개선을 얻을 수 있다. 둘째, 방법의 재사용성이다. CBHEM의 기여는 특정 모델의 내부 구조가 아니라 학습 데이터의 구성 층위에 있으므로, 향후 더 발전된 주행 모델로 교체하더라도 큐레이션 절차는 그대로 적용된다. 셋째, 점진적 무인화의 근거 축적이다. 고정 노선을 반복 운행하는 버스의 특성상 동일 구간의 약점을 지속적으로 진단·보강할 수 있으며, 취약 구간의 개선이 통계적으로 확인되는 만큼 안전관리원 동승 단계에서 완전 무인 단계로 나아가는 판단의 정량적 근거가 된다.

활용 측면에서 본 방법론은 판교 판타G버스와 같은 자율주행 셔틀을 비롯하여, 심야버스·공단 순환버스 등 고정 노선을 반복 운행하는 대중교통 전반에 적용될 수 있다. 운행에서 축적되는 실패 기록이 곧 다음 학습의 입력이 되는 폐루프 구조는, 노선이 늘어날수록 개선이 누적되는 확장 가능한 운용 체계를 제공한다. 이를 통해 운전 인력 부족과 교통 소외 지역 문제에 대응하는 지속 가능한 자율주행 대중교통의 토대를 마련할 수 있을 것으로 기대한다.

참고문헌

[1] Bojarski, Mariusz et al. (2016), "End to End Learning for Self-Driving Cars", arXiv preprint arXiv:1604.07316.

[2] Ross, Stéphane, Geoffrey J. Gordon&Drew Bagnell (2011), "A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning", Proceedings of the 14th International Conference on Artificial Intelligence and Statistics: 627-635.

[3] Zhou, Weitao et al. (2025), "DRARL: Disengagement-Reason-Augmented Reinforcement Learning for Efficient Improvement of Autonomous Driving Policy", Robotics and Autonomous Systems 193: 105059.

[4] Liu, Deqing et al. (2025), "TakeAD: Preference-Based Post-Optimization for End-to-End Autonomous Driving with Expert Takeover Data", IEEE Robotics and Automation Letters.

[5] Yang, Yanding et al. (2026), "DTCCL: Disengagement-Triggered Contrastive Continual Learning for Autonomous Bus Planners", Proceedings of the IEEE Intelligent Vehicles Symposium.

[6] Shrivastava, Abhinav, Abhinav Gupta&Ross Girshick (2016), "Training Region-Based Object Detectors with Online Hard Example Mining", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition: 761-769.

[7] Cui, Yin et al. (2019), "Class-Balanced Loss Based on Effective Number of Samples", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition: 9268-9277.

[8] Codevilla, Felipe et al. (2019), "Exploring the Limitations of Behavior Cloning for Autonomous Driving", Proceedings of the IEEE/CVF International Conference on Computer Vision: 9329-9338.

[9] Dosovitskiy, Alexey et al. (2017), "CARLA: An Open Urban Driving Simulator", Proceedings of the 1st Annual Conference on Robot Learning: 1-16.

[10] International Organization for Standardization (2018), Intelligent Transport Systems — Adaptive Cruise Control Systems — Performance Requirements and Test Procedures, ISO 15622:2018, Geneva: ISO.

[11] 차세대융합기술연구원 (2024), 판타G버스 자율주행 주행 데이터(상태정보·제어정보·주행데이터·GPS/INS·객체인식), 공공데이터 개방 자료.