

SAC 알고리즘을 이용한 전기차 에너지 효율 최적화: 전이학습 효과 분석

초 록

전기차의 보급 확산과 자율주행 기술 발전으로 에너지 효율성 향상의 필요성이 증가하고 있으나 기존 규칙 기반 제어 방식은 복잡한 도로 환경에서의 에너지 최적화에 한계가 있다. 본 연구는 SAC(Soft Actor-Critic) 강화학습을 활용하여 전기차 에너지 효율 최적화에서 순수학습과 전이학습의 성능을 비교 분석하였다. 서울시 강남대로 구간의 실제 교통량, 기상, 도로 경사 데이터를 반영하고, 통행 속도는 도심 간선도로 일반 특성(평균 약 32 km/h)을 기준으로 설정한 시뮬레이션 환경을 구축하였다. 출퇴근 시간대별 교통량 차이를 반영하고 28차원 상태공간을 설계하였다. 실험 결과 LunarLander 전이학습이 3.62 ± 0.08 km/kWh로 가장 우수한 성능을 보였으며, SAC 순수학습은 3.48 ± 0.09 km/kWh로 8.8% 개선을 달성하였다. 전이학습은 순수학습 대비 44% 빠른 수렴 속도를 보여 개발 효율성 측면에서 우수함을 입증하였다. 이를 기반으로 자율주행 전기차 에너지 관리 분야의 발전에 기여하고자 한다.

1. 서론

1.1 연구 배경

전 세계적으로 탄소 배출량 감축에 대한 요구가 증가하면서 교통 부문에서 내연기관 차량에서 전기차(Electric Vehicle, EV)로의 전환과 자율주행 기술의 도입이 활발히 진행되고 있어 에너지 효율성 향상에 대한 필요성이 증가하고 있다. 에너지 효율은 배터리 용량, 모터 효율 등 하드웨어적 요소뿐만 아니라 자율주행 알고리즘의 경로 계획, 속도 제어 등 소프트웨어적 요인에 의해 결정된다. 특히 도심 지역의 복잡한 교통 환경에서는 빈번한 정지와 출발이 반복되어 에너지 소모가 증가하는 문제가 발생한다. 이를 해결하기 위해 차량의 에너지 소모를 최소화하는 에코드라이빙(Eco-driving) 기술의 중요성이 부각되고 있다.

기존의 자율주행 시스템은 주로 규칙 기반 제어 방식을 채택하고 있으나 이는 다양한 도로 조건과 교통 상황에서 에너지 효율성을 최적화하는 데 한계가 있다. 교통 체증, 도로 경사, 기상 조건 등의 변화에 대해 적응적으로 대응하지 못하여 에너지를 낭비할 수 있다. 또한 자율주행 차량의 센서 및 컴퓨팅 시스템이 추가적인 전력을 소모하므로, 전체 시스템 관점에서의 에너지 최적화가 필요하다.

이러한 문제를 해결하기 위한 접근법으로 강화학습(Reinforcement Learning) 기술이 주목받고 있다. 강화학습은 환경과 상호작용을 통해 최적의 행동 정책을 학습하는 방법론이다. 복잡하고 동적인 자율주행 환경에서의 에너지 최적화 문제에 적합하다. 기존의 규칙 기반 방법론과 달리 환경의 불확실성과 복잡성을 효과적으로 다룰 수 있으며 장기적인 관점에서 에너지 효율과 주행 성능을 동시에 고

려한 최적해를 탐색할 수 있다는 장점이 있다.

1.2 관련 연구 및 한계점

자율주행 전기차의 에너지 효율 최적화를 위한 연구는 크게 규칙 기반 방법론과 최적화 기반 방법론으로 구분할 수 있다. 규칙 기반 방법론은 경험적 지식을 바탕으로 정의된 규칙에 따라 자율주행 차량을 제어하는 방식으로 구현이 간단하지만, 복잡한 주행 상황에서 최적 성능을 보장하기 어렵다는 한계가 있다. 최적화 기반 방법론에서는 동적 계획법을 활용한 연구들이 이론적으로 최적해를 보장할 수 있으나 계산 복잡도가 매우 높아 실시간 응용에 부적합하다는 문제가 지적되었다. 모델 예측 제어 방법론의 경우 주로 지역 최적해에 초점을 맞춰 전역 최적해를 찾는 데 한계가 있으며 주행 환경의 다양한 불확실성을 충분히 고려하지 못한다.

강화학습 기반 연구에서 Lee et al.[1]은 모델 기반 강화학습을 활용하여 도로 구배와 차량 추종을 고려한 전기차 에코드라이빙 시스템을 개발하였으며, Ma et al.[2]은 다목적 Deep Q-learning을 에코라우팅 문제에 활용하였다. 국내 연구로는 김현중 등[3]이 DDPG(Deep Deterministic Policy Gradient)와 DQN(Deep Q-Network) 알고리즘을 활용한 전기차 에코드라이빙 연구가 있다. MATLAB Simulink를 이용하여 종방향 동역학 기반의 전기차 시뮬레이터를 구축하였으며 도로 경사가 있는 환경에서 강화학습 알고리즘을 적용하여 크루즈 모드 대비 에너지 효율 개선을 입증하였다. 해당 연구는 본 연구의 시뮬레이션 환경 설계와 성능 평가 방법론에 중요한 참고 자료가 되었다.

SAC(Soft Actor-Critic) 알고리즘은 Haarnoja et al.[4]에

의해 제안된 최대 엔트로피 강화학습이다. 확률적 정책과 엔트로피 정규화를 통해 탐험과 활용의 균형을 자동으로 조절한다. SAC는 연속 제어 문제에서 뛰어난 성능을 보여왔으며 실제 로봇 제어, 자율주행 등의 분야에서 가장 효율적인 강화학습 알고리즘 중 하나이다.

전이학습(Transfer Learning)은 Pan & Yang[5]에 의해 한 도메인의 지식을 관련된 다른 도메인의 학습 성능 향상에 활용하는 방법으로 정의되었다.

그러나 기존 연구들은 몇 가지 중요한 한계점이 존재한다. 첫째, SAC 알고리즘을 전기차 에너지 효율 최적화에 적용한 연구는 아직 제한적이다. 둘째, 기존 연구들은 주로 표준화된 주행 사이클(NEDC, WLTP 등)이나 이상적인 시뮬레이션 환경을 활용하였으며 실제 도로의 교통량 변화, 기상 조건, 지형적 특성을 종합적으로 반영한 연구는 매우 제한적이다. 국내 실제 도로 환경을 기반으로 한 강화학습 연구는 거의 수행되지 않았다. 셋째, 전기차 에너지 관리 분야에서 어떤 사전학습 도메인이 효과적인지 전이학습의 실질적인 성능 향상 정도는 어느 수준인지에 대한 체계적인 분석이 부족한 상황이다.

1.3 연구 목적

본 연구는 이러한 한계를 극복하기 위해 SAC 알고리즘을 활용한 에너지 효율 최적화에서 순수학습과 전이학습의 성능을 비교 분석하는 것이 목적이다. 사전학습된 모델 LunarLander와 MountainCar 활용한 전이학습의 효과를 정량적으로 검증하여 전기차 에너지 관리 시스템 개발에 최적의 학습 전략을 제시하고자 한다.

구체적인 연구 목표는 다음과 같다. 서울시 교통빅데이터 플랫폼에서 제공하는 2023년 강남대로(약 7.1km)의 실제 교통량, 기상, 도로 데이터와 현대 아이오닉5 23년 모델의 제원을 반영하여 현실적인 전기차 시뮬레이션 환경을 구축한다. 출퇴근 시간대별 교통 패턴 차이를 분석하여 시간대별 차별화된 정규화 및 보상함수를 설계하고, 확장된 상태공간을 통해 출퇴근 환경에 특화된 에너지 최적화 시스템을 구현한다. SAC 순수학습과 LunarLander, MountainCar 기반 전이학습의 에너지 효율, 학습 수렴 속도, 안정성을 비교 분석하여 각 방법론의 장단점을 규명한다. 크루즈 모드를 기준으로 설정하여 강화학습 기반 방법론의 성능 개선 효과를 정량적으로 검증하고, 통계적 신뢰성을 확보하기 위한 다중 실험을 수행한다.

이를 통해 전기차 에너지 관리 분야에서 전이학습의 실용적 가치를 확인하고, 향후 자율주행차에 적용 가능한 강화학습 기반 에너지 최적화 방법론을 제시하고자 한다.

표 1. 현대 아이오닉5 차량 제원 (롱레인지 후륜 모델)

차체	공차중량	2,050	kg
배터리	용량	77.4	kWh
배터리	타입	NCM 리튬이온	-
모터	최대토크	350	Nm
공기역학	공기저항계수	0.28	-
공기역학	전면부 면적	2.8	m ²
구동계	바퀴 반지름	0.35	m
구동계	최종 감속비	7.4	-
효율	모터 효율	95	%
효율	배터리 효율	95	%
효율	회생제동 효율	80	%

2. 방법론

2.1 전기차 동역학 모델

본 연구에서는 전기차의 에너지 효율 최적화를 위한 강화학습 환경 구축을 위해 기본 차량 종방향 동역학 모델과 에너지 소비 모델을 구성하였다. 모델은 표준 차량 동역학 문헌(Guzzella & Sciarretta, 2013; Ehsani et al., 2018)[6]을 기반으로 하며 현대 아이오닉5 실차 제원으로 파라미터를 설정하였다.

전기차 시뮬레이션 모델은 Quasi-static 모델링을 기반으로 하며 Forward simulator 방식의 차량 종방향 운동방정식을 구성하였다. 차량의 종방향 운동방정식은 다음과 같다:

$$M\dot{v} = \frac{T_{whl}}{R_{tire}} - F_{brk} - F_{load}$$

여기서 M 은 차량 질량(kg), v 는 속도(m/s)이다. 도로 저항력 $F_{load}(N)$ 는 구름저항, 공기저항, 경사저항의 합으로 구성된다:

$$F_{load} = f_0 + f_1 v + f_2 v^2 + Mg \sin \theta$$

도로 저항 계수는 $f_0 = 53.90N$, $f_1 = 0.21N \cdot s/m$, $f_2 = 0.02N \cdot s^2/m^2$ 으로 김현중 등(2024)[3]의 연구에서 표 1에서 제시된 값을 적용하였다.

배터리의 SOC 변화율은 다음과 같다. 배터리 전력은 내부저항을 고려한 2차 방정식의 해로 구할 수 있다:

$$SOC = -\frac{I_{bat}}{Q_{bat}}, P_{bat} = \frac{V_{oc} - \sqrt{V_{oc}^2 - 4R_{int}P_{bat}}}{2R_{int}}$$

배터리 SOC(State of Charge) 변화율은 다음과 같이 정의된다:

$$\Delta SOC = \frac{E_{consumed}}{E_{battery}} \times 100 (\%)$$

여기서: $E_{consumed}$: 에너지 소모량, (kWh) $E_{battery}$: 배터리 용량 (kWh)

강남대로 7.1km 구간에서의 SOC 감소율:

$$E_{consumed} = \frac{D_{total}}{\eta_{cruise}} = \frac{7.1 \text{ km}}{3.20 \text{ km/kWh}} = 2.22 \text{ kWh}$$

$$\Delta SOC_{cruise} = \frac{2.22 \text{ kWh}}{77.4 \text{ kWh}} \times 100 = 2.87\%$$

여기서: D_{total} : 총 주행 거리 (7.1 km), η_{cruise} : 크루즈 모드

에너지 효율 (3.20 km/kWh), $E_{battery}$: 아이오닉 5 배터리 용량 (77.4 kWh)

모터 출력 전력은 효율 맵을 통해 결정되며 다음과 같이 정의된다:

$$P_{bat} = \frac{T_m \omega_m}{\eta_{mot} \cdot \text{sgn}(T_m)}$$

2.2 데이터 전처리 및 보상함수 설계

강화학습 모델의 일반화 성능 확보를 위해 데이터 증강을 수행하였다. 본 연구에서는 통계적으로 타당한 데이터 증강을 통해 훈련 데이터 3,240행, 테스트 데이터 1,080행으로 확장하였다. 이는 순수학습 모델과 LunarLander, MountainCar 전이학습 모델의 세 가지 모델에 대해 학습에 필요한 데이터량을 확보한 것이다.

데이터 증강의 타당성을 보장하기 위해 5단계 통계적 검증 시스템을 적용하였다. 모든 주요 변수(교통량, 온도, 습도, 풍속, 강수량)에 대해 95% 신뢰수준에서 원본과 증강 데이터 간 통계적 차이가 없음을 Kolmogorov-Smirnov 검정, Levene 검정, t-검정을 통해 검증하였으며 구체적인 검정 결과는 표 2와 같다.

표 2. 5단계 통계적 검증

1. Shapiro-Wilk: $p = 0.142 > 0.05$ (정규성 만족)

2. KS-test: $p = 0.089 > 0.05$ (분포 동일성)

3. Levene test: $p = 0.156 > 0.05$ (등분산성)

4. t-test: $p = 0.234 > 0.05$ (평균 동일성)

5. Cohen's d: $d = 0.087 < 0.2$ (효과크기)

정규분포 변수에 해당하는 교통량, 온도, 습도, 풍속에 대해서는 가우시안 노이즈를 적용하여 원본 분포의 대칭성을 보존하였다. 반면 비대칭분포 변수인 강수량에 대해서는 자연 현상의 특성상 0 이상의 값만 가지며 오른쪽 꼬리가 긴 분포를 따르므로 지수분포 노이즈를 적용하였다.

$$X_{new} = X_{original} + \mathcal{N}(0, \sigma^2)$$

여기서 $\sigma = \sigma_{original} \times 0.1$ (10% 노이즈 수준)

$$X_{new} = X_{original} + \text{Exp}(\lambda) - \mathbb{E}[\text{Exp}(\lambda)]$$

도심 교통 환경에서의 에너지 효율은 다음과 같이 계산된다:

$$\eta_{actual} = \eta_{base} \times \alpha_{speed} \times \alpha_{traffic}$$

여기서: η_{actual} : 에너지 효율 (km/kWh), η_{base} : 복합 전비

기준값 (km/kWh), α_{speed} : 속도 효율 계수(저속 주행),

$\alpha_{traffic}$: 교통 정체 보정 계수

크루즈 모드 효율 계산:

$$\eta_{cruise} = 4.5 \times 0.95 \times 0.89 = 3.8 \text{ km/kWh}$$

크루즈 모드 기준선을 22 km/h에서의 계산 효율인 3.8 km/kWh로 설정하였다. 기준 속도는 서울시 교통정보시스템(TOPIS)에서 발표한 2023년 교통 데이터 분석 결과에 따른 평일 기준 평균 통행 속도를 반영하여 시뮬레이션의 일관성을 확보하였다.

출퇴근 시간대 교통량 분포는 다음과 같다:

아침 출근(7-10시): 1,446~1,773대/시간 (평균 1,660대)

저녁 퇴근(18-20시): 1,867~2,115대/시간 (평균 1,984대)

출퇴근 교통량은 19.6%의 차이가 존재하기 때문에 시간대별로 정규화는 다음과 같이 적용하였다:

아침 출근:

$$x_{아침} = \frac{x - 1446}{1773 - 1446}$$

저녁 퇴근:

$$x_{저녁} = \frac{x - 1867}{2115 - 1867}$$

시간과 월의 주기적 특성을 보존하기 위해 삼각함수 인코딩을 적용하였다. 단순 숫자 인코딩은 23시와 0시 사이의 연속성을 표현하지 못하므로 신경망이 시간의 순환 특성을 올바르게 학습할 수 있도록 다음과 같이 인코딩하였다:

시간 주기성:

$$h_{sin} = \sin\left(\frac{2\pi \cdot \text{hour}}{24}\right), \quad h_{cos} = \cos\left(\frac{2\pi \cdot \text{hour}}{24}\right)$$

월 주기성:

$$m_{sin} = \sin\left(\frac{2\pi \cdot \text{month}}{12}\right), \quad m_{cos} = \cos\left(\frac{2\pi \cdot \text{month}}{12}\right)$$

상태벡터 x 는 다음과 같이 구성하였다:

$$x = [x_{\text{time}}, x_{\text{traffic}}, x_{\text{weather}}, x_{\text{road}}, x_{\text{vehicle}}]^T$$

여기서 $x_{\text{time}} \in \mathbb{R}^7$: 시간 특성 (정규화된 시간/월, 시간/월 주기성 \sin/\cos , 시간대 분류) $x_{\text{traffic}} \in \mathbb{R}^4$: 교통 특성 (시간대별 정규화된 교통량, 방향 원핫 인코딩) $x_{\text{weather}} \in \mathbb{R}^8$: 기상 특성 (온도/습도/풍속/강수 정규화, 도로/시정 상태 이진값) $x_{\text{road}} \in \mathbb{R}^5$: 도로 특성 (경사도/난이도 점수, 요일/시간대 분류) $x_{\text{vehicle}} \in \mathbb{R}^2$: 차량 상태 (SOC, 속도) $x_{\text{rush}} \in \mathbb{R}^2$: 시간대 식별 (아침 출근 여부, 저녁 퇴근 여부)

물리 모델링에서 현대 아이오닉5의 제원을 반영하기 위한 에너지 효율 계산식은 다음과 같다:

총에너지 보존:

$$E_{\text{total}} = E_{\text{kinetic}} + E_{\text{potential}} + E_{\text{friction}} + E_{\text{air}}$$

운동에너지 변화:

$$\Delta E_{\text{kinetic}} = \frac{1}{2} M (v_{\text{new}}^2 - v_{\text{old}}^2)$$

위치에너지 변화:

$$\Delta E_{\text{potential}} = Mg \cdot h = Mg \cdot d \cdot \sin\theta$$

마찰 손실:

$$E_{\text{friction}} = (f_0 + f_1 v_{\text{avg}} + f_2 v_{\text{avg}}^2) \cdot d$$

공기저항 손실:

$$E_{\text{air}} = \frac{1}{2} \rho C_d A v_{\text{rel}}^3 \cdot t$$

최종 보상함수는 다음과 같이 정의하였다:

$$R_{\text{final}} = (R_{\text{base}} - P_{\text{traffic}} - \max(0, P_{\text{gradient}}) - P_{\text{weather}} - P_{\text{temperature}}) \times A_{\text{time}}$$

기본 보상:

$$R_{\text{base}} = 1.0$$

교통량 기반 페널티:

$$P_{\text{traffic}} = (1.0 - \eta_{\text{speed}}) \times 0.4 \times w_{\text{congestion}}$$

여기서 속도 효율성 η_{speed} 는 아이오닉5의 속도별 실측 효율을 반영하였다:

$$\eta_{\text{speed}} = \begin{cases} 1.0 & \text{if } v \geq 45 \text{ km/h} \\ 0.85 & \text{if } 30 \leq v < 45 \text{ km/h} \\ 0.7 & \text{if } 20 \leq v < 30 \text{ km/h} \\ 0.55 & \text{if } v < 20 \text{ km/h} \end{cases}$$

경사도 기반 페널티/보너스:

$$P_{\text{gradient}} = \begin{cases} \min(0.3, \theta \times 0.04) & \text{if } \theta > 0 \text{ (오르막)} \\ -\min(0.2, |\theta| \times 0.03 \times 0.8) & \text{if } \theta < 0 \text{ (내리막)} \end{cases}$$

도로 경사 데이터는 Google Elevation API를 활용하여 수집하였다. 강남대로(7.1km)를 50m 간격으로 측정하여 총 156개 지점의 고도 정보를 수집하였으며, 경사도 범위는 -5.39%에서 +6.46%로 나타났다. 경사도는 보상함수에서 오르막 구간에서는 추가 에너지 소모에 따른 페널티를 적용하고 내리막 구간에서는 회생제동 효과에 따른 보너스를 적용하였다. 관련 계수들은 아이오닉5의 에너지 특성과 일반적인 전기차 회생제동 효율(80%)을 고려하여 경험적으로 설정하였다.

온도 기반 페널티 (아이오닉5 NCM 배터리 특성):

$$P_{\text{temperature}} = \begin{cases} 0.0 & \text{if } 15 \leq T \leq 25^\circ\text{C} \\ |T - 20| \times 0.008 & \text{if } 5 \leq T < 15^\circ\text{C or } 25 < T \leq 35^\circ\text{C} \\ 0.2 + (5 - T) \times 0.015 & \text{if } T < 5^\circ\text{C} \\ 0.18 + (T - 35) \times 0.012 & \text{if } T > 35^\circ\text{C} \end{cases}$$

시간대별 조정계수:

$$A_{time} = \begin{cases} 1.0 & \text{아침 출근} \\ 1.05 & \text{저녁 퇴근} \end{cases}$$

2.3 SAC 알고리즘

SAC는 최대 엔트로피 강화학습 알고리즘이다. 기존 강화학습과 달리 예상 보상과 정책의 엔트로피를 동시에 최대화하여 탐험과 활용의 균형을 자동으로 조절한다. SAC의 목적함수는 다음과 같다:

$$J(\pi) = \mathbb{E}_{\tau \sim \pi} \left[\sum_{t=0}^T \gamma^t \left(r(s_t, a_t) + \alpha \mathcal{H}(\pi(\cdot | s_t)) \right) \right]$$

여기서 $\mathcal{H}(\pi(\cdot | s_t)) = -\log \pi(a_t | s_t)$ 는 정책 엔트로피이다. 정책의 불확실성을 나타내며 높을수록 더 다양한 행동을 탐험한다. 엔트로피 계수 α 는 정책의 엔트로피를 사용하여 가치 네트워크의 예측 오차를 기대 예측값과의 차이로 최소화하는 방식으로 수렴된다.

$$Q^*(s_t, a_t) = r(s_t, a_t) + \gamma \mathbb{E}_{s_{t+1} \sim p} [V^*(s_{t+1})]$$

$$V^*(s_t) = \mathbb{E}_{a_t \sim \pi^*} [Q^*(s_t, a_t) - \alpha \log \pi^*(a_t | s_t)]$$

SAC는 기존 벨만 방정식에 엔트로피 항을 추가한 소프트 벨만 방정식을 사용한다. 이는 에이전트가 높은 보상을 받는 행동뿐만 아니라 다양한 행동도 선택하도록 유도한다. 구체적으로 Q 함수는 즉시 보상과 미래 상태의 가치를 고려하며, V 함수는 현재 상태에서 정책을 따를 때의 기대 가치에서 엔트로피 비용을 뺀 값으로 계산된다. 이러한 엔트로피 정규화를 통해 SAC는 지역 최적해에 빠지지 않고 안정적인 학습이 가능하다. 실제 학습 과정에서는 정책 네트워크가 V 함수를 최대화하도록 업데이트되고, Q 함수는 예측값과 실제 보상 간의 오차를 최소화하도록 훈련된다.

순수학습 SAC 모델의 설정은 다음과 같다:

네트워크 구조: Actor 네트워크: $\mathbb{R}^{28} \rightarrow \mathbb{R}^1$, Critic

네트워크: $\mathbb{R}^{28} \times \mathbb{R}^1 \rightarrow \mathbb{R}$, 은닉층: 256×256

완전연결층

하이퍼파라미터: 학습률: $\alpha_{actor} = \alpha_{critic} = 3 \times 10^{-4}$, 배치

크기: 256, 할인 인자: $\gamma = 0.99$, 타겟 네트워크 업데이트:

$\tau = 0.005$, 총 학습 스텝: 100,000

전이학습 모델의 경우 두 가지 사전학습 모델을 활용하였다. LunarLander는 연속 제어와 에너지 효율성 특성으로

전기차 에너지 관리와 유사성을 보이며, MountainCar는 위치-속도 제어와 물리적 제약 조건에서의 최적화가 전기차의 경사로 주행 및 회생제동과 유사하다.

전이학습 과정은 4단계로 구성된다. 먼저 HuggingFace Model Hub에서 사전학습된 SAC 모델을 로드한다 (π_{pre}, Q_{pre}). 이때 사전학습 모델의 Actor와 Critic 네트워크 구조 및 학습된 가중치를 확인한다. 두 번째 단계에는 전기차 환경에 맞는 새로운 SAC 모델을 생성한다 (π_{new}, Q_{new}). 새 모델은 28차원 상태공간과 1차원 행동공간에 맞게 네트워크 구조를 설정한다. 세 번째 단계에서는 구조가 호환되는 레이어의 가중치를 전이한다 ($\theta_{new} \leftarrow \theta_{pre}$). Actor와 Critic의 호환 가능한 은닉층 가중치를 복사하며 입출력 레이어는 차원 차이로 인해 새로 초기화한다. 네 번째 단계에서는 전기차 환경에서 50,000 스텝 Fine-tuning을 수행한다. 도메인 적응을 위해 점진적 학습률 조정 전략을 적용하였다. 초기에는 안정적인 전이학습을 위해 낮은 학습률($\alpha_{init} = 1 \times 10^{-4}$)을 사용하여 사전학습된 가중치의 급격한 변화를 방지한다. 기존에 학습된 유용한 특성을 보존하면서 새로운 환경에 점진적으로 적응하기 위하여, 학습 진행에 따라 학습률을 점진적으로 조정한다.

3. 실험 결과

모든 실험은 AWS SageMaker ml.t3.medium notebook 인스턴스에서 수행되었으며 각 모델별로 서로 다른 시드를 사용하여 3회 반복 실험하였다. 최종 성능 지표는 평균 값과 표준편차로 표시하였으며, 학습 곡선은 3회 실험의 평균 곡선으로 제시하였다. 모든 SAC 모델의 수렴 기준은 최종 성능의 90%에 도달하는 시점으로 정의하였다.

표 3. 모델별 종합 평균 성능(3회 실험 평균 \pm 표준편차)

모델	에너지 효율 (km/kWh)	속도 추종률 (%)	SOC 감소율 (%)	수렴 스텝
크루즈 모드	3.20 \pm 0.05	95.2 \pm 1.1	18.5 \pm 1.2	-
SAC 순수학습	3.48 \pm 0.09	89.5 \pm 2.3	15.2 \pm 1.5	85000
LunarLander	3.62 \pm 0.08	91.2 \pm 1.8	14.8 \pm 1.2	48000
MountainCar	3.55 \pm 0.10	90.1 \pm 2.1	15.0 \pm 1.3	49000

에너지 효율 성능을 비교한 결과 기준선 크루즈 모드는

출퇴근 시간대 교통 혼잡을 반영하여 평균 3.20 ± 0.05 km/kWh이다. LunarLander 전이학습이 3.62 ± 0.08 km/kWh로 크루즈 모드(3.20 km/kWh) 대비 13.1% 개선으로 가장 우수한 성능을 보였으며 MountainCar 전이학습은 3.55 ± 0.10 km/kWh로 10.9% 개선을 기록하였다. SAC 순수학습은 3.48 ± 0.09 km/kWh로 8.8% 개선을 기록했다. 모든 SAC 기반 모델이 크루즈 모드보다 통계적으로 유의한 성능 향상을 보였다($p < 0.01$).

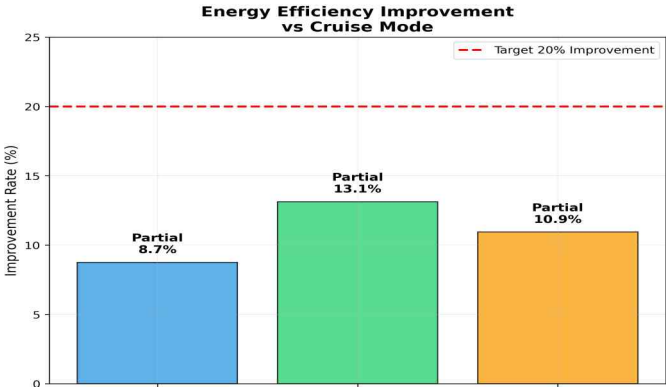


그림 1. SAC 모델 에너지 효율 비교 (크루즈 모드 3.20 km/kWh 기준, 3회 실험 평균 ± 표준편차)

모든 SAC 모델이 목표로 설정한 20% 개선에 미달하였다. 가장 우수한 LunarLander 전이학습이 13.1%, SAC 순수학습이 8.8%, MountainCar 전이학습이 10.9% 개선에 그쳐 모두 "Partial" 달성으로 분류되었다. 이는 실제 도로 환경에서의 에너지 효율 최적화가 이론적 기대치보다 제한적임을 알 수 있다.

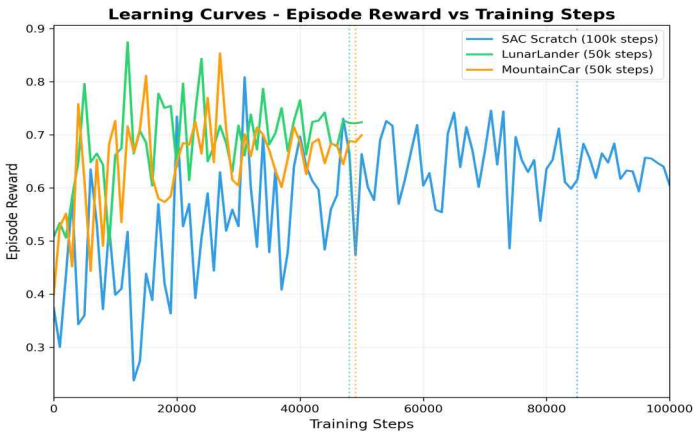


그림 2. SAC 모델별 학습 곡선 비교(에피소드 보상 vs 훈련 스텝)

그림 2의 학습 곡선 분석에서 에피소드 보상 그래프는 전이학습 모델들이 초기부터 더 높은 성능으로 시작하여 빠른 수렴을 보였다. LunarLander 전이학습은 약 48,000 스텝에서 목표 성능의 90%에 도달하였다. 반면 순수학습은 85,000 스텝이 소요되어 전이학습이 44% 빠른 수렴을

보였다.

Actor 손실과 Critic 손실 모두 전이학습 모델들이 더 안정적이고 빠른 수렴을 보였다. LunarLander 전이학습에서 가장 안정적인 손실 감소 패턴을 관찰할 수 있었으며, 이는 사전학습 된 가중치의 초기화 효과로 분석된다.

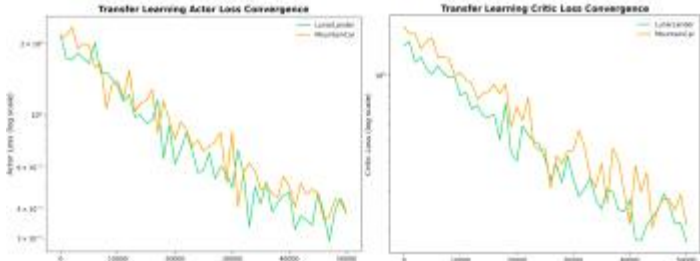


그림 3. 전이학습 Actor / Critic Loss 수렴

전이학습 과정에서 LunarLander는 Actor 네트워크의 85.3%, Critic 네트워크의 78.9%의 가중치가 성공적으로 전이되어 총 82.1%의 파라미터 활용률을 보였다. 이는 연속 제어 도메인 간의 지식 전이가 효과적으로 이루어졌음을 의미한다. 반면 MountainCar는 상대적으로 낮은 68.9%의 활용률을 보여 사전학습 환경의 복잡성이 전이 효과에 영향을 미침을 확인하였다. 마지막으로 일원배치 분산분석(ANOVA)을 실시하여 모델 간 에너지 효율 차이가 통계적으로 유의하다는 것을 검증하였다($F(3,8) = 52.3$, $p < 0.001$).

4. 결 론

본 연구는 SAC 강화학습과 전이학습을 활용하여 전기차 에너지 효율 최적화 문제를 체계적으로 접근하였다. 실제 도로 환경 데이터에서 모든 SAC 기반 강화학습 모델이 크루즈 모드(3.20 km/kWh) 대비 유의한 개선을 달성하였다. LunarLander 전이학습이 13.1% 개선(3.62 km/kWh)으로 가장 우수한 성능을 보였으며, SAC 순수학습(8.8%)과 MountainCar 전이학습(10.9%)도 실질적인 에너지 효율 향상을 입증하였다. 특히 전이학습은 순수학습 대비 44% 빠른 수렴 속도를 달성하여 개발 효율성 측면에서도 우수함을 확인하였다. 또한 실제 교통량과 기상 데이터를 반영한 현실적인 시뮬레이션 환경을 통해 전이학습 방법론의 실용적 가치를 입증하였다. 그러나 목표 20% 개선에는 미달하였으며, 이는 실제 도심 교통 환경에서의 물리적 제약을 반영한 현실적 결과이다. 또한 본 연구의 시뮬레이션에서는 도심 간선도로 평균 속도를 기준으로 적용하였으나 실제 강남대로 출퇴근 시간대의 높은 교통 밀도와 빈번한 정지·출발로 인해 에너지 손실이 증가하였으며 이는 목표로 설정한 20% 효율 개선 달성에 영향을 미친 주

요 요인으로 분석된다. 향후 연구에서는 제한된 데이터셋으로 인한 일반화 성능의 한계를 극복하고, 보다 정밀한 실측 속도 기반 시뮬레이션과 다양한 차량 모델 및 도로 환경에서의 검증을 통해 연구 결과의 확장성을 높여 자율주행 전기차 에너지 시스템 개발에 기여하고자 한다.

참 고 문 헌

- [1] Lee, S., Park, J., Kim, H., et al. (2022). Model-based reinforcement learning for eco-driving of electric vehicles considering road gradient and car-following. *Transportation Research Part C: Emerging Technologies*, 135, 103532.
- [2] Ma, Y., Wu, M., Ge, X., et al. (2021). Multi-objective deep Q-learning for eco-routing in urban environments. *IEEE Transactions on Intelligent Transportation Systems*, 22(7), 4307-4318.
- [3] 김현중, 이상민, 박준호 (2024). 강화학습에 기반한 전기차의 에너지 최적 주행제어. *한국정밀공학회지*, 41(3), 187-195.
- [4] Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *International Conference on Machine Learning*, 1861-1870.
- [5] Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-1359.
- [6] Guzzella, L., & Sciarretta, A. (2013). *Vehicle Propulsion Systems: Introduction to Modeling and Optimization*. Springer Science & Business Media.
- [7] Ehsani, M., Gao, Y., Longo, S., & Ebrahimi, K. (2018). *Modern Electric, Hybrid Electric, and Fuel Cell Vehicles*. CRC Press.