

Multi-Task Learning을 활용한 PVT v2 프레임워크 성능 개선 - 시퀀스 다이어그램

8조

202001156 정보통계학과 김수영

202002510 컴퓨터융합학부 송재현

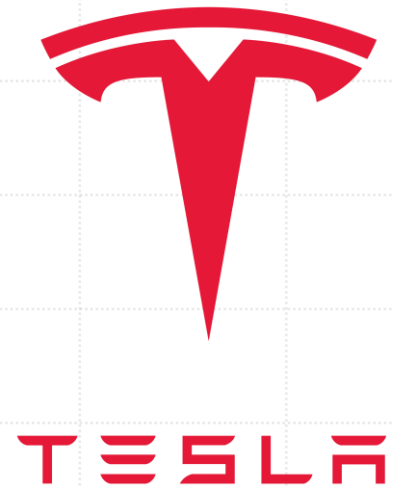


목차

- 연구 배경
- 연구 목적
- 연구 질문/가설
- 소프트웨어 사용 사례 / DIAGRAM
- 문제 해결에 대한 사용 사례 / DIAGRAM
- 해결 방법에 대한 알고리즘 순서도

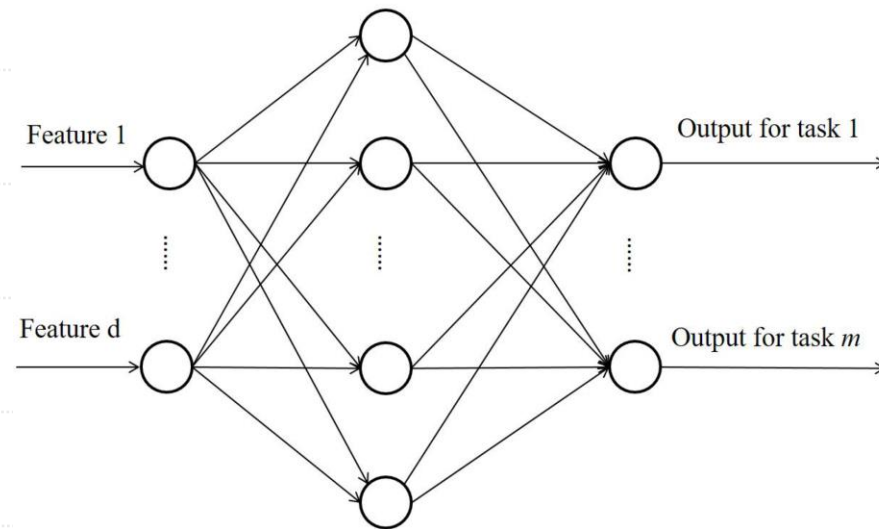
연구 배경

- MTL은 향후 AI 성능 발전에 큰 영향을 줄 수 있는 패러다임
- 일반화, 경량화에 유리
- 테슬라에서는 자율주행기술에 MTL을 적용시키는 연구를 진행 중



연구 배경

- Multi-Layer feedforward neural network: 원시적인 형태의 MTL 적용 모형
- 초기 MTL 모형은 여러 출력 노드 사용



연구 배경

- MTL 전용 구조들도 활발히 연구 중
 - MuT, M3ViT, IPT 등 → 다양한 태스크를 효율적으로 처리 가능
- 최근 연구에서는 STL 모델을 MTL로 확장하는 방식의 효율성 입증
 - SwinMTL → Swin Transformer 기반 MTL 확장으로 **메모리 사용량 감소 및 정확도 유지**
- 강력한 STL 모델인 PVT v2를 MTL로 확장 시 성능 향상 기대

연구 배경

- 기존 강력한 STL 모델을 기준으로 삼아 MTL을 적용, 그 효과를 정량적으로 평가하는 방식이 현재 유효한 접근법
- 본 연구는 PVT v2 기반 MTL 모델을 설계하여 STL 대비 성능 향상과 자율주행 분야 적용 가능성을 분석하고자 함

연구 목적

- PVT v2를 MTL 기반으로 확장하여 성능 향상 여부를 검증
- 자율주행 환경에서 요구되는 이미지 분류, 객체 탐지, 의미론적 분할 작업을 하나의 모델로 처리
- Swin MTL 등의 기존 모델보다 성능 개선 가능성 탐색
- 경량화, 확장성, 정확도 향상을 통해 실제 적용 가능성 제고

연구 질문/가설

- RQ1.
PVT v2 프레임워크에 MTL을 적용한 모형은 단일 작업 학습 모형, 다른 MTL 모형에 비해 작업(이미지 분류, 객체 탐지, 의미론적 분할)에 대한 유의미한 성능 향상이 이루어 지는가?
- RQ2.
MTL을 이용해 학습시킨 모형은 자율주행 분야에서 기존 프레임워크 대비 어떤 장단점을 갖는가?

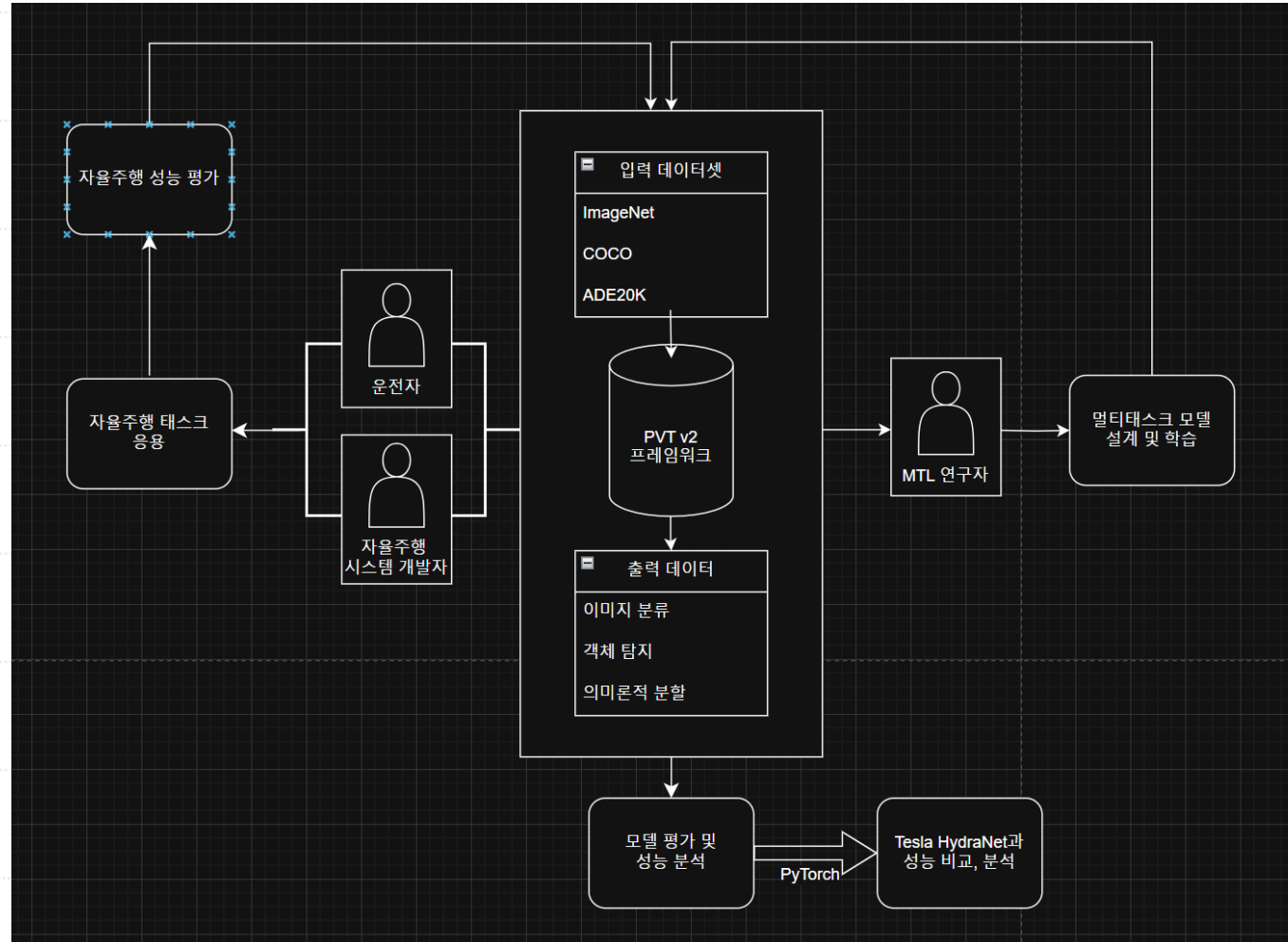
연구 질문/가설

- H1.
MTL을 이용해 학습시킨 PVT v2 모형은 단일 작업 모형, 다른 MTL 모형보다 정확도, AP, mIoU, #Param 등의 성능 지표에서 유의미한 개선을 보일 것이다.
- H2.
MTL 기반 모형은 자율 주행 분야에서 연구되는 모형인 HydraNet 대비 정확도 측면에서 유의미한 성능 향상을 보여줄 것이다.

소프트웨어 사용 사례 / DIAGRAM

주요 Actor	자율주행 시스템 개발자, 운전자, MTL 연구자
주요 기능 구성 요소	<p>주요 기능</p> <ul style="list-style-type: none">- 자율주행 태스크 응용- 객체 탐지, 의미론적 분할, 이미지 분류- 멀티태스크 학습 모델 설계 및 학습- 모델 평가 및 성능 분석 <p>구성 요소</p> <ul style="list-style-type: none">- PVT v2 프레임워크 및 MTL을 위한 디코더
입/출력 데이터	<p>입력데이터(결과):이미지데이터셋 (ImageNet, COCO,ADE20K)</p> <p>출력 데이터(결과): 이미지 분류 레이블, 객체 탐지 박스, 의미론적 분할 마스크</p>
데이터 Flow	<ol style="list-style-type: none">1) 이미지 데이터 전처리2) MTL 모델 학습3) 테스트 및 성능 분석
외부 시스템 연계	<p>모델 비교 대상: Tesla HydraNet</p> <p>평가 툴: PyTorch, sklearn을 활용</p> <p>추가적인 모형 활용 시:Hugging Face</p>

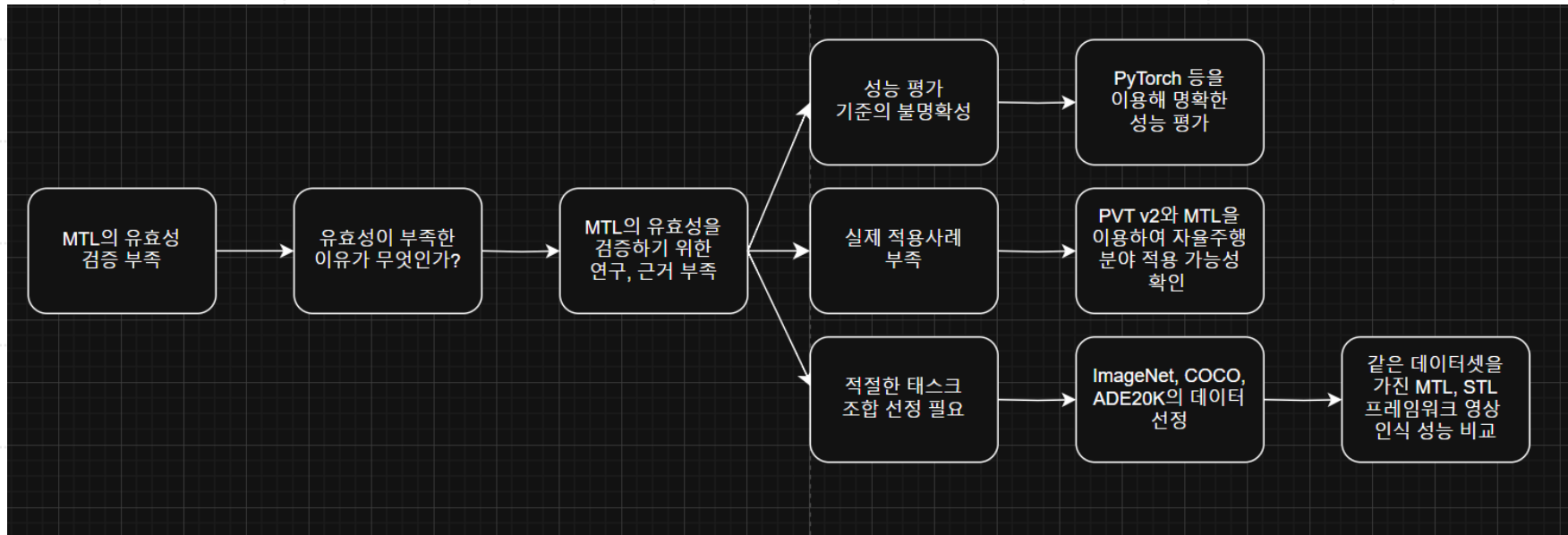
소프트웨어 사용 사례 / DIAGRAM



문제 해결에 대한 사용 사례 / DIAGRAM

핵심 문제	MTL의 실질적인 유효성 검증 부족
직접 요인	MTL의 유효성에 대한 연구와 근거 부족
간접 요인	실제 적용사례 부족, 적절한 태스크 조합의 선정 및 성능 평가 기준의 불명확성
활용 맥락	자율주행시스템, 영상인식기반 응용분야

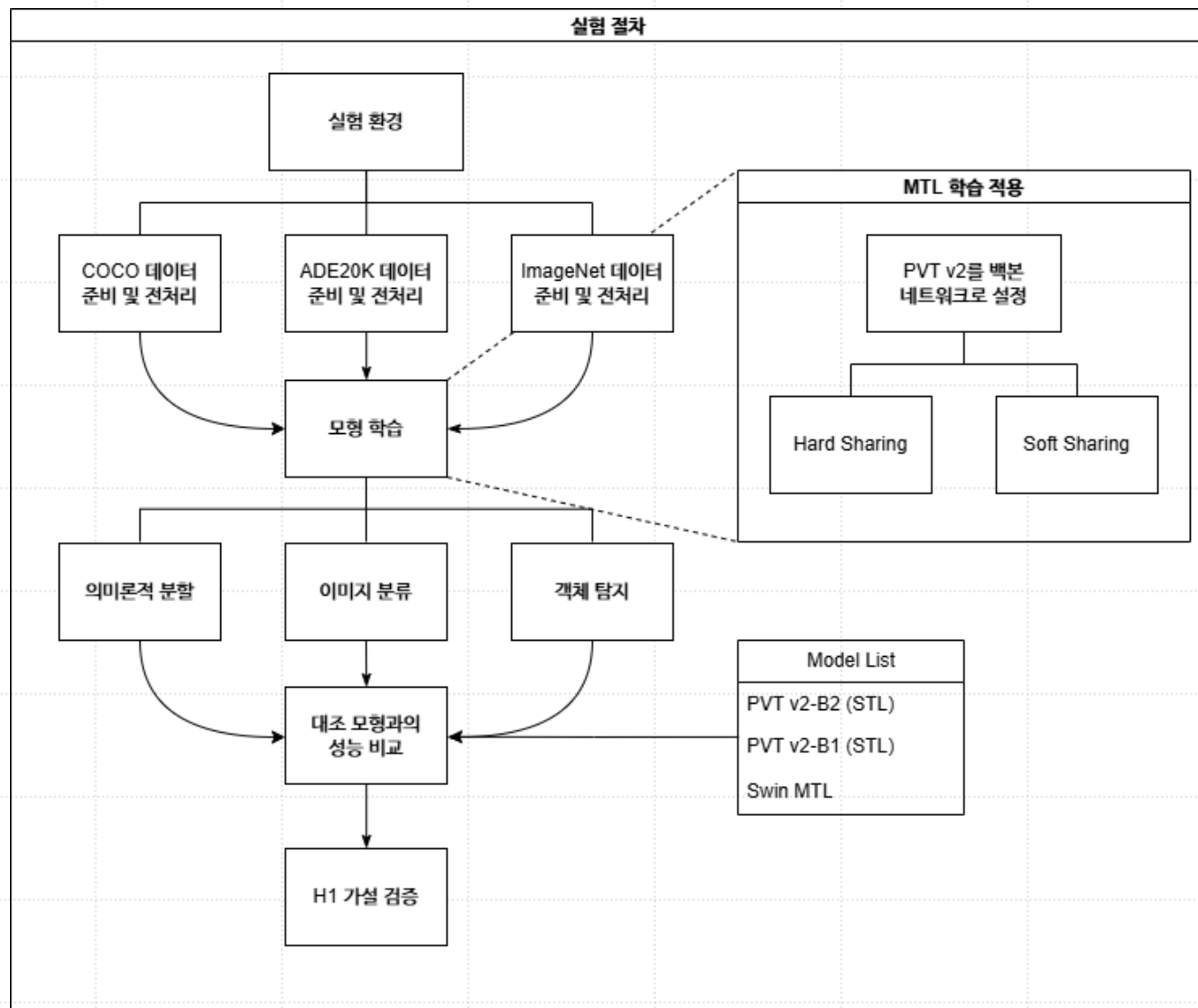
문제 해결에 대한 사용 사례 / DIAGRAM



해결 방법에 대한 알고리즘 순서도

실험 환경	COCO, ADE20K, ImageNet 데이터 준비 및 전처리
모형 학습	전처리된 데이터로 의미론적 분할, 이미지 분류, 객체 탐지 학습
대조 모형과의 성능 비교	<ul style="list-style-type: none">• PVT v2-B2 (STL)• PVY v2-B1 (STL)• Swin MTL
H1 가설 검증	MTL을 이용해 학습시킨 PVT v2 모형은 단일 작업 모형보다 정확도, AP, mIoU, #Param 등의 성능 지표에서 유의미한 개선을 보일 것이다.

해결 방법에 대한 알고리즘 순서도



AI 도구 활용 정보

사용 도구	GPT-4.5
사용 목적	논문 탐색
프롬프트	<ul style="list-style-type: none">● MTL 기반 Transformer를 연구한 논문들을 찾아줘● 기존 STL 기반 모형을 MTL로 확장시킨 연구를 진행한 논문을 찾아줘
반영 위치	<ol style="list-style-type: none">1. 인터뷰 질문 목록 (p.5)2. 아이디어 설명 문단 정리 (p.6)
수작업 수정	없음