

Multi-Task Learning을 활용한 PVT v2 프레임워크 성능 개선 - 테스트 계획 및 디자인

8조

202001156 정보통계학과 김수영

202002510 컴퓨터융합학부 송재현

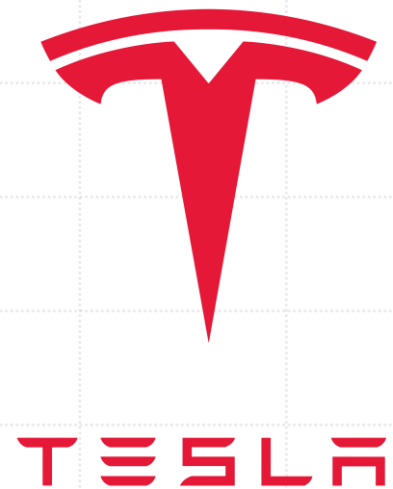


목차

- 연구 배경
- 연구 질문/가설
- TEST PLAN
- TEST CASES

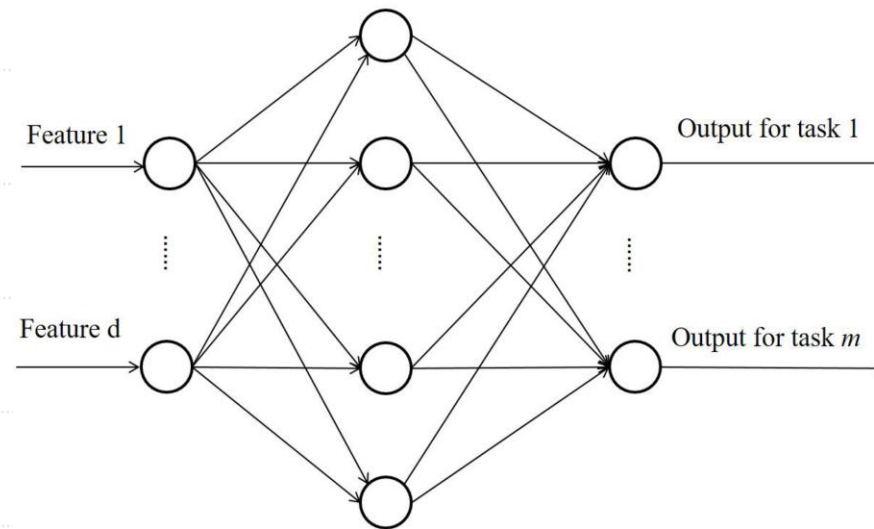
연구 배경

- 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
 - MulT, 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 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.

테스트 상세

- 독립 변수
 - 학습 구조 (STL vs MTL, Hard Sharing, Soft Sharing)
 - 백본 모델 종류
- 종속 변수
 - 성능 지표: Accuracy, AP (Average Precision), mIoU (mean Intersection over Union), #Parameters, 연산 속도



테스트 상세

- 데이터셋
 - ImageNet(224*224) or ImageNet(32*32)
 - COCO 2017
 - ADE20K 2016
- 사용 장비 및 도구
 - GPU: NVIDIA GTX 30080 10GB
 - 프레임워크: PyTorch
 - 모델: PVTv2 기반 구조 + MTL 확장(Hard Sharing, Soft Sharing)

테스트 상세

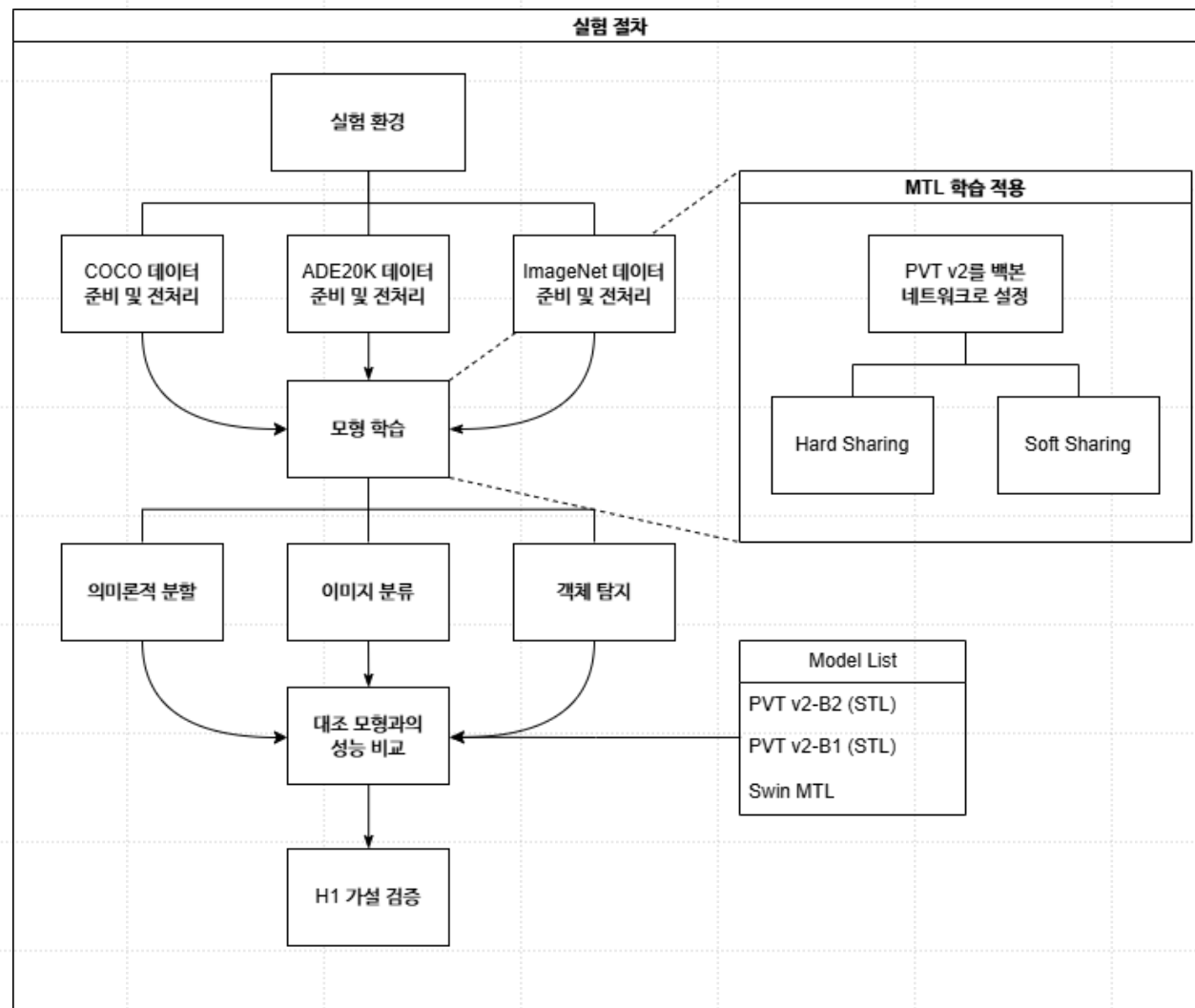
■ 실험 흐름

1. 데이터셋 준비 및 전처리
2. PVTv2 모델을 MTL 학습 구조로 확장
3. 세 가지 작업에 대해 학습
4. 기준 모델(Swin MTL, PVTv1-B1)과 비교
5. 지표를 기반으로 유의미함 검증

■ 정량 지표

- Accuracy(이미지 분류)
- AP (객체 탐지)
- mIoU (의미론적 분할)
- #Parameters, Inference Time

- 활용 도구: COCO API, mIoU 평가 스크립트



테스트 케이스

Id	대상(모델/조건)	실험 조건	테스트 데이터	평가지표	예상 결과
TC-1.1	PVT v2-B0 (STL)	단일 작업 학습	ImageNet	Accuracy	Accuracy 70%
TC-1.2	PVT v2-B0 (STL)	단일 작업 학습	COCO	AP	AP 37.1
TC-1.3	PVT v2-B0 (STL)	단일 작업 학습	ADE20K	mIoU	mIoU 37.2%
TC-2	Swin MTL	MTL 구조 (기존 구조)	3개 데이터셋	Accuracy, AP, mIoU	Acc: 81.3% AP: 50.5 mIoU: 50.35%
TC-3	PVT v2-B0 + MTL (Hard Sharing)	다중 작업, 공유 구조	3개 데이터셋	Accuracy, AP, mIoU	Acc: 75% AP: 43 mIoU: 43%



테스트 케이스

- Accuracy: 이미지 분류 정확도
- AP: 객체 탐지에서의 평균 정밀도
- mlou: 의미론적 분할에서의 평균 IoU (Intersection over Union, 교집합 비율)
- #Param: 모델의 전체 파라미터 수
- Inference Time: 테스트 1장당 소요 시간(실시간 분석용)