##### Project Document

Project Brainstorming Result

|  |  |
| --- | --- |
| Project Name | Multi-Task Learning의 최적 공유 전략 연구 확장 및 실험적 유효성 분석 |

8조

202001156 김수영

202002510 송재현

지도교수: 이종률 교수님 (서명)

Document Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| Rev# | Date | Affected Section | Author |
| 1 | 2023/03/26 | PVT v2: Improved Baselines with Pyramid Vision Transformer | W. Wang et al. |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Contents

[1. (문제 해결을 위한) 아이디어 발산 5](#_Toc193878719)

[2. 아이디어 수렴 5](#_Toc193878720)

[3. 시각화하기 6](#_Toc193878721)

[4. AI 도구 활용 정보 8](#_Toc193878722)

List of Figure

**그림 목차 항목을 찾을 수 없습니다.**

# (문제 해결을 위한) 아이디어 발산

본 연구에서는 Pyramid Vision Transformer v2(PVT v2)를 기반으로 Multi-Task Learning(MTL)에서의 최적 공유 전략을 탐색하고 실험적으로 검증하고자 한다. PVT v2는 기존 Pyramid Vision Transformer v1(PVT v1) 프레임워크를 개선하여 새로운 비전 Transformer 백본 네트워크를 제시한 것으로, 주요 개선 사항으로는 선형 공간 축소 어텐션(Linear Spatial Reduction Attention, Linear SRA), 겹치는 패치 입베딩(Overlapping Patch Embedding), 합성곱 기반 FFN(Convolutional Feed-Forward Network, CFFN)이 있다. PVT v2는 기존 모델 대비 연산 효율성이 높으며, 다양한 컴퓨터 비전 작업에서 우수한 성능을 보인다. 해당 모형을 활용하여 이미지 분류, 객체 탐지, 그리고 의미론적 분할(Semantic Segmentation)을 데이터셋으로 MTL 학습을 수행하고, MTL을 사용하지 않는 기존 모형과의 성능 비교, MTL의 성능 향상을 검증한다.

연구 키워드 정리

* Multi-Task Learning
* Pyramid Vision Transformer v2: 연구에서 사용할 백본 모형
* Image Classification: MTL 학습에 사용할 데이터셋
* Object Detection: MTL 학습에 사용할 데이터셋
* Semantic Segmentation: MTL 학습에 사용할 데이터셋

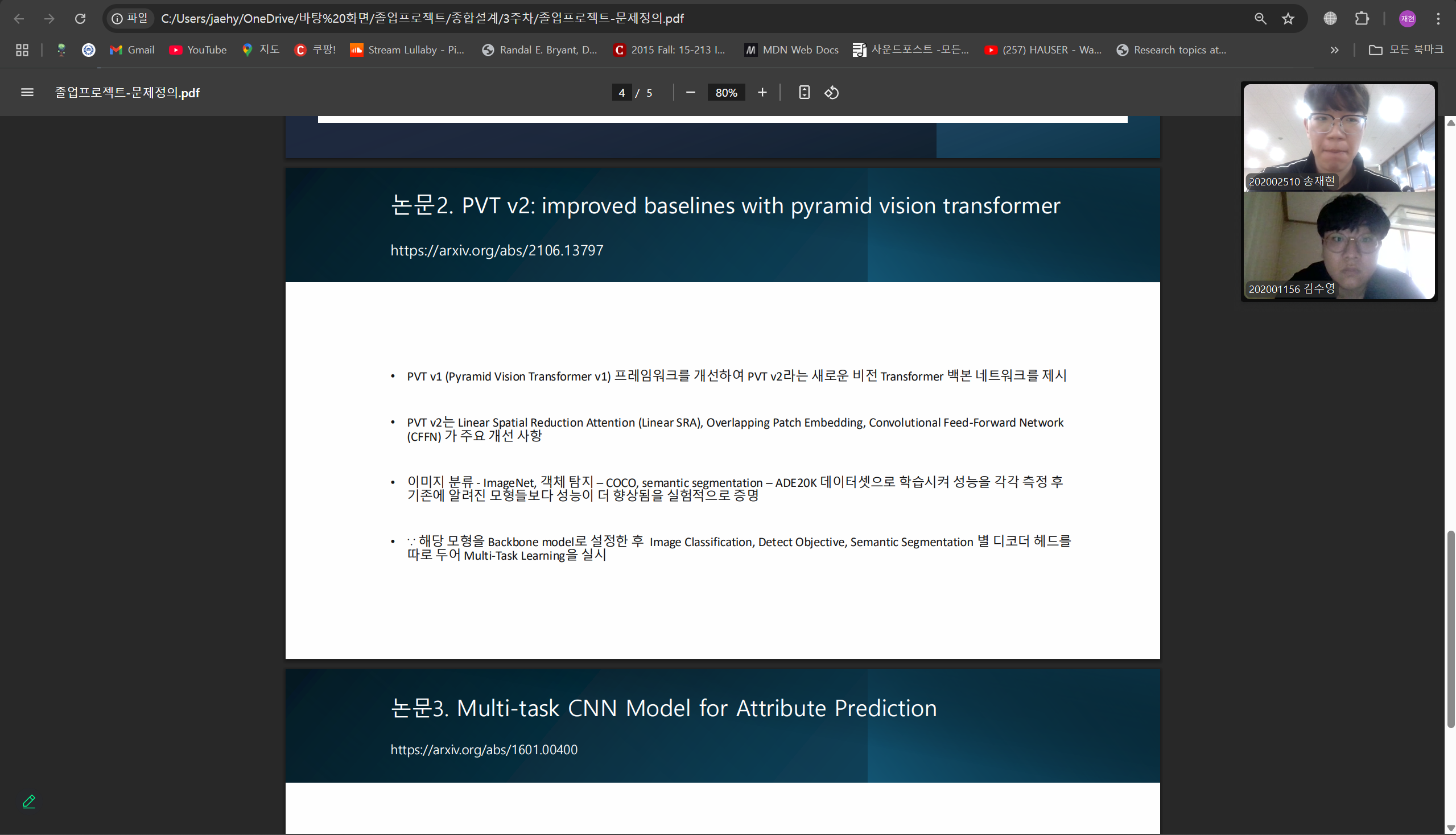
주요 이론 및 연구 변수

* 독립변수(IV)

1. Backbone의 공유 전략: 완전 공유, 부분 공유, 개별 Backbone 사용
2. Decoder의 적용 방식: Task-Specific Decoder

* 종속변수(DV)

1. 이미지 분류 성능
2. 객체 탐지 성능
3. 의미론적 분할 성능



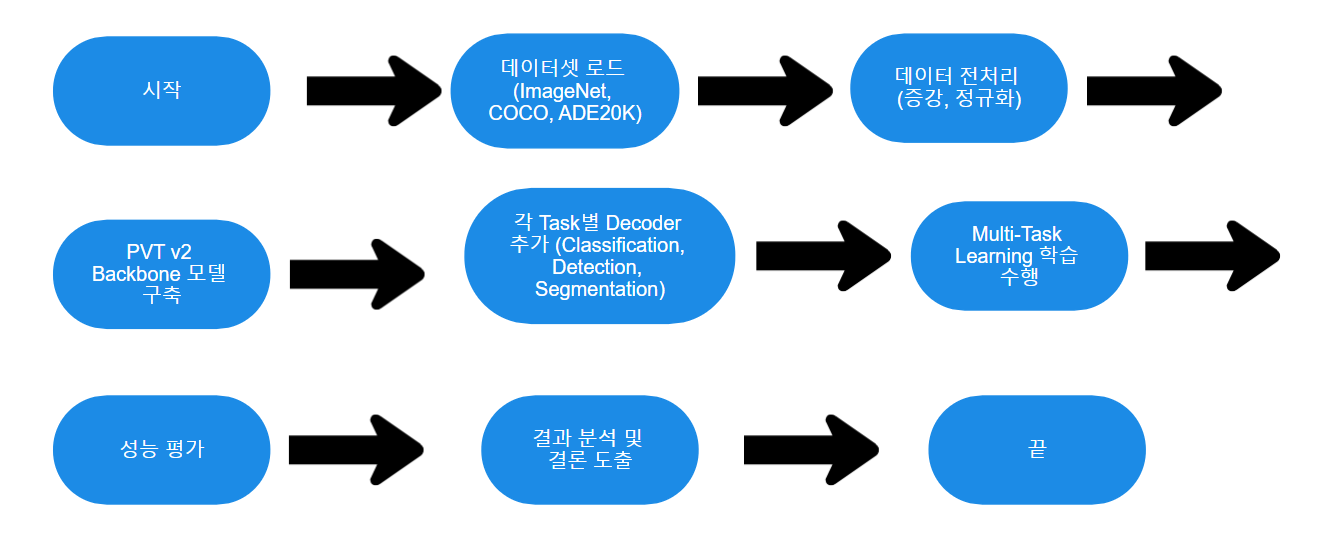
아이디어 발산 회의 장면 캡쳐

# 아이디어 수렴

|  |  |
| --- | --- |
| 핵심 개념 | 정의 및 하위 개념 목록 |
| Multi-Task Learning | 1. 하나의 모델이 여러 개의 컴퓨터 비전 작업을 동시에 수행하는 학습 방법  2. 하나의 모델이 한 개의 컴퓨터 비전 작업을 수행하는 Single-Task Learning과 비교 |
| PVT v2 Backbone | 본 연구에서 백본으로 설정하여 MTL 학습 작업을 수행할 모델 구조 |
| 데이터셋 | 본 연구에서 MTL 학습을 수행할 데이터로, ImageNet(이미지 분류), COCO(객체 탐지), ADE20K (의미론적 분할)이 있다 |
| 성능 평가 방법 | 모델의 정확도, 연산의 효율성, 연산량 |
| Task-Specific Decoders | 각 작업에 특화된 Decoder을 활용, 최적의 성능 도출 목적 |

# 시각화하기

연구 일정에 대한 Flow chart



# AI 도구 활용 정보

|  |  |
| --- | --- |
| 사용 도구 | GPT-4 |
| 사용 목적 | 연구 계획의 방향성에 대한 질문 |
| 프롬프트 | * 내가 지금 Multi-Task Learning에 대한 주제를 바탕으로 졸업프로 젝트를 진행중이야. 학부생 수준에서 10주 안에 수행해야하기에 굳이 어렵게 하지 않으려 해. 현재 구체적인 연구 계획은 정해지지 않았고 연구 방향성은 Multi-Task Learning의 유효성 검증에 관한 것을 바 탕으로 할 예정이야. 그래서 연구를 어떻게 진행해야할지 구체적인 문제 정의를 하고 싶어. 일단 내가 원하는 방향성은 기존에 알려져 있 는 모형들을 Multi-Task Learning으로 학습시켜 본 후 성능향상이 실제로 이루어지는지, Multi-Task Learning의 실질적인 유효성을 검증을 보여주는 것이야. 이러한 방향성을 바탕으로 연구 목표를 더 구체화해줘. 예를 들어 최신 모형을 바탕으로 Multi-Task Learning 을 수행한다 할 시 써먹을 논문을 구체적으로 정한다거나, 실제 이용 되고 있는 AI서비스의 모형들을 Multi-Task Learning으로 수행해보 는 등 말이야. * 1. 컴퓨터비전과 관련된 태스크를 염두에 두고 있어 2. 기본적인 성능 기준은 정확도(Accuracy)야 하지만, 만약 성능이 비슷하거나 약간 안좋게 나온다고 하여도 파라미터의 수가 유의미하 게 줄었으면 그것 또한 경량화가 잘 됐다는 증거로써 성능에 포함시 킬 예정이야. 3. 컴퓨팅자원과 데이터셋에 대한 제약은 없어 4. 잘 알려진 모델을 사용하는 것이 좋겠지? 실제 상용화된 모형이 Multi-Task Learning으로 학습될 경우 그에 대한 성능이 실제로 향 상되는 것이 목적이니까 |
| 반영 위치 | 1. 인터뷰 질문 목록 (p.5) 2. 아이디어 설명 문단 정리 (p.6) |
| 수작업  수정 | 있음(논리 보강, 사례 교체 등) |