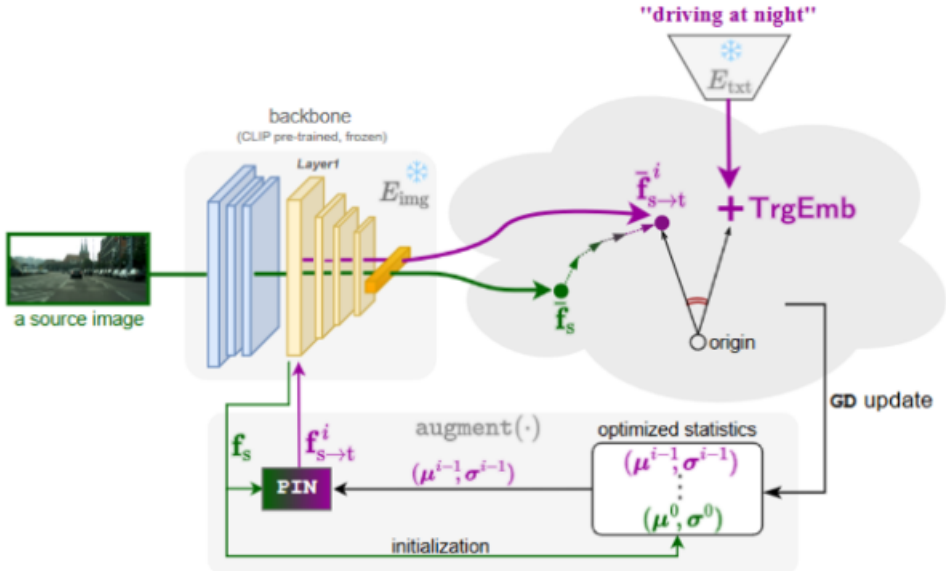




## 2024-하계 집중이수제 주간학습보고서 (2주차)

창의과제	Unsupervised domain adaption 이미지 Segmentation 알고리즘 연구				
이름	서민정	학습기간	2024.07.01. ~ 2024.07.07		
학번	21011591	학습주차	2주차	학습시간	12
학과(전공)	컴퓨터공학과	과목명	자기주도창의전공 II	수강학점	3
※ 수강학점에 따른 회차별 학습시간 및 10회차 이상 학습 준수					
금주 학습목표	PØDA: Prompt-driven Zero-shot Domain Adaptation 논문 리뷰 및 문제 정의				
학습내용	<p><b>[PØDA: Prompt-driven Zero-shot Domain Adaptation 논문 리뷰]</b></p> <p>Domain adaptation은 훈련 시점에서 target domain의 데이터를 필요로 한다. 그러나 Computer Vision 분야에서 적절한 이미지를 충분한 양으로 구하는 것은 어려운 과제이다.</p> <p>이에 본 논문에서는 target domain의 이미지를 전혀 사용하지 않고 zero-shot domain adaptation을 하는 것을 목표로 한다. 이를 위해 CLIP 모델의 contrast learning 기반 latent space를 활용하여 img domain을 text domain으로써 사용한다.</p> <p>논문의 방법론은 아래와 같다.</p>  <p>[그림 1] PØDA의 CLIP latent space 활용법</p> <p><b>1. Source Domain 임베딩 추출</b></p> <ul style="list-style-type: none"> <li>- 중간 layer의 output인 low-level feature <math>f_s</math>와 CLIP 임베딩 <math>f_s</math> 추출</li> </ul>				



	<p><b>2. 스타일 주입</b></p> <ul style="list-style-type: none"> <li>- target text imbedding과 source img imbedding이 가까워질 수 있는 <math>\sigma</math>와 <math>\mu</math>를 추출하고, 이 수와 기존 <math>f_s</math>를 PIN의 input으로 하여 <math>f_s \rightarrow t</math> 생성</li> </ul> $PIN(f_s, \mu, \sigma) = \sigma \left( \frac{f_s - \mu(f_s)}{\sigma(f_s)} \right) + \mu,$ <p><b>3. 손실 계산 및 스타일 update</b></p> <ul style="list-style-type: none"> <li>- 기존 <math>f_s</math>와 <math>f_s \rightarrow t</math>의 코사인 거리를 손실함수로 활용하여 이를 최소화하도록 스타일 (<math>\mu, \sigma</math>) 업데이트</li> </ul> <p><b>4. 해당 <math>\mu</math>와 <math>\sigma</math>를 활용하여 text prompt만으로 unseen domain의 img 임베딩 추출</b></p> $f_s^t = PIN(f_s, \mu_t, \sigma_t)$ <p><b>5. 기존 모델에 해당 PIN을 붙여 활용</b></p> <p><math>M' = (M_{feat}, M'_{cls})</math>  text prompt에서 제시한 domain의 특화 모델이 됨</p> <p><b>[문제 정의 및 가설 설정]</b>  논문의 주된 목표인 “target domain의 이미지를 전혀 사용하지 않고 zero-shot domain adaptation을 하는 것”을 해치지 않으면서 본 모델의 성능을 개선해보고자 한다.</p> <p>이를 위해 아래 세 개의 가설을 제시하였다.</p> <ul style="list-style-type: none"> <li>- PODA에서의 text driven 방식은 Domain gap이 존재하므로, optimization시 synthesis image를 활용하여 domain gap을 줄인다.</li> <li>- text feature에 노이즈를 준다면 좀 더 다양한 domain을 표현할 수 있을 것이다.</li> <li>- text feature와 synthesis image를 통한 image feature를 동시에 사용하여 style transfer를 위한 변수들을 optimize 한다면 둘의 장점을 모두 가져올 수 있을 것이다.</li> </ul>
학습방법	<p><b>[개별 논문 리뷰 및 노선 공유 페이지를 통한 정리]</b>  개별 학습을 통하여 PODA를 이해하는 시간을 가졌다. 이때, 공유 노선 페이지에 정리하여 협업 및 스터디에 도움이 되도록 하였다.</p> <p><b>[세미나를 통한 가설 소개 및 구현 가능성 검토]</b>  모임을 가지며 개별 학습한 내용을 토대로 가설을 공유하는 시간을 가졌다. 한명씩 자신이 생각한 가설을 발표하였고, 이에 질문하며 구현 가능성을 검토하는 시간을 가졌다.</p>
학습성과 및 목표달성도	<p>100%</p> <p>PODA의 논문을 리뷰하고 inference하는 시간을 가져보며, 구조와 방법론을 이해할 수 있었다. 또한 세미나를 통해 각자 생각해본 가설을 공유하고, 구현 가능성을 검토하여 앞으로의 방향성을 성공적으로 수립할 수 있었다.</p>



참고자료 및 문헌	<a href="https://arxiv.org/abs/2212.03241">https://arxiv.org/abs/2212.03241</a> <a href="https://github.com/astra-vision/PODA">https://github.com/astra-vision/PODA</a>
내주 계획	GPU 사용을 위한 GCP 환경 세팅 및 코드 reimplementation 진행

2024 년 07 월 07 일

지도교수

김세원

(인)