

# Seminar Feedback Report (Week10)

2021312738 소프트웨어학과 김서환

10주차 세미나의 내용은 이미소 연구원의 발표로 시작되었으며, 주제는 다중 장면 절대 자세 회귀(Multi-Scene Absolute Pose Regression) 작업에서의 트랜스포머 모델의 자기 주의(Self-Attention) 메커니즘 활성화 문제를 다루는 연구였다. 첫 번째로, 연구팀은 기존 트랜스포머 기반 위치 추정 모델에서 자기 주의가 충분히 활성화되지 않는 원인으로 쿼리-키 임베딩 공간의 왜곡과 위치 임베딩의 비효율성을 지적하였다. 이에 대한 해결책으로는 쿼리-키 정렬 손실(Query-Key Alignment Loss)과 고정 위치 인코딩(Fixed Positional Encoding)을 제안하였고, 이를 통해 자기 주의 메커니즘의 표현력이 향상되었으며, 다양한 장면에서의 위치 추정 성능도 기존 방법보다 우수함을 입증하였다. 실험 결과, 주의 맵의 엔트로피가 증가하며 모델이 보다 복잡한 관계를 학습할 수 있게 되었음을 보여주었고, 이번 발표는 트랜스포머 기반 컴퓨터 비전 모델의 구조적 한계를 개선하려는 시도와 그 실질적인 성과를 확인할 수 있어 의미 깊다고 생각했다.

그 다음은 이민규 연구원의 발표로, 주제는 ESOP(Autoencoded Supervision for Perceptual Image Super-Resolution)이었다. 이 연구는 CVPR 2025에 채택될 만큼 주목받는 내용으로, AI 기반 슈퍼 해상도(Super Resolution, SR) 모델이 직면하고 있는 지각 품질과 왜곡의 상충(perception-distortion tradeoff) 문제를 해결하고자 하였다. 기존의 SR 모델들은 고해상도를 복원할 때 충실도와 현실감 사이에서 균형을 맞추기 어려운 점이 있었는데, 이에 비해 ESOP는 오토인코더를 활용한 지도 신호(Autoencoded Supervision)를 통해 충실도 편향(fidelity bias)만을 효과적으로 최소화하고, 동시에 시각적 품질(perceptual quality)을 유지하도록 설계되었다. 실험 결과, 다양한 데이터셋에서 기존 SOTA(superior of the art) 기법보다 우수한 성능을 보였고, 특히 섬세한 디테일 복원에서 향상된 결과를 도출하였다. 이번 발표는 이미지 복원 분야에서 모델의 지각 품질과 수치적 정확도를 동시에 고려한 혁신적 접근을 보여주었다는 점에서 매우 인상 깊었다.