

연구 배경



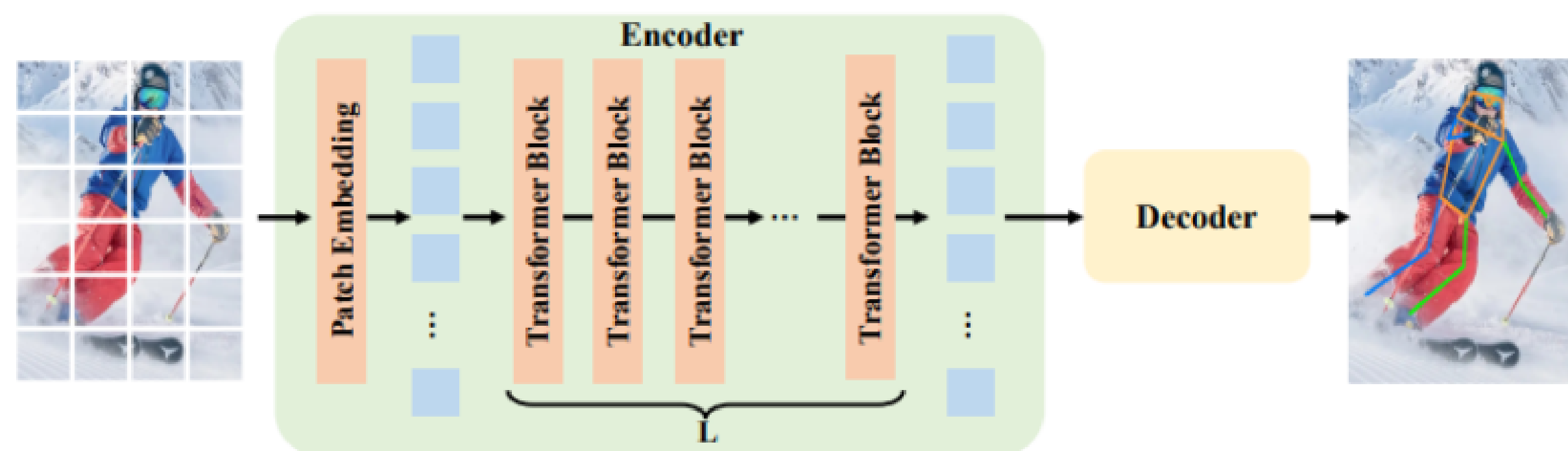
사진을 찍을 때, 누구나 더 자연스럽고 멋진 자세로 나오고 싶어 합니다. 하지만 스스로 포즈를 잡다 보면 어딘가 어색하게 느껴질 때가 많고, 이런 순간에 누군가 옆에서 적절한 조언을 해준다면 얼마나 좋을까 하는 생각이 들었습니다.

이러한 아이디어를 구체화하기 위해, "자세를 추천하고 수정해주는 머신러닝 모델을 만들어보자"라는 목표로 자기주도 프로젝트를 시작했습니다.

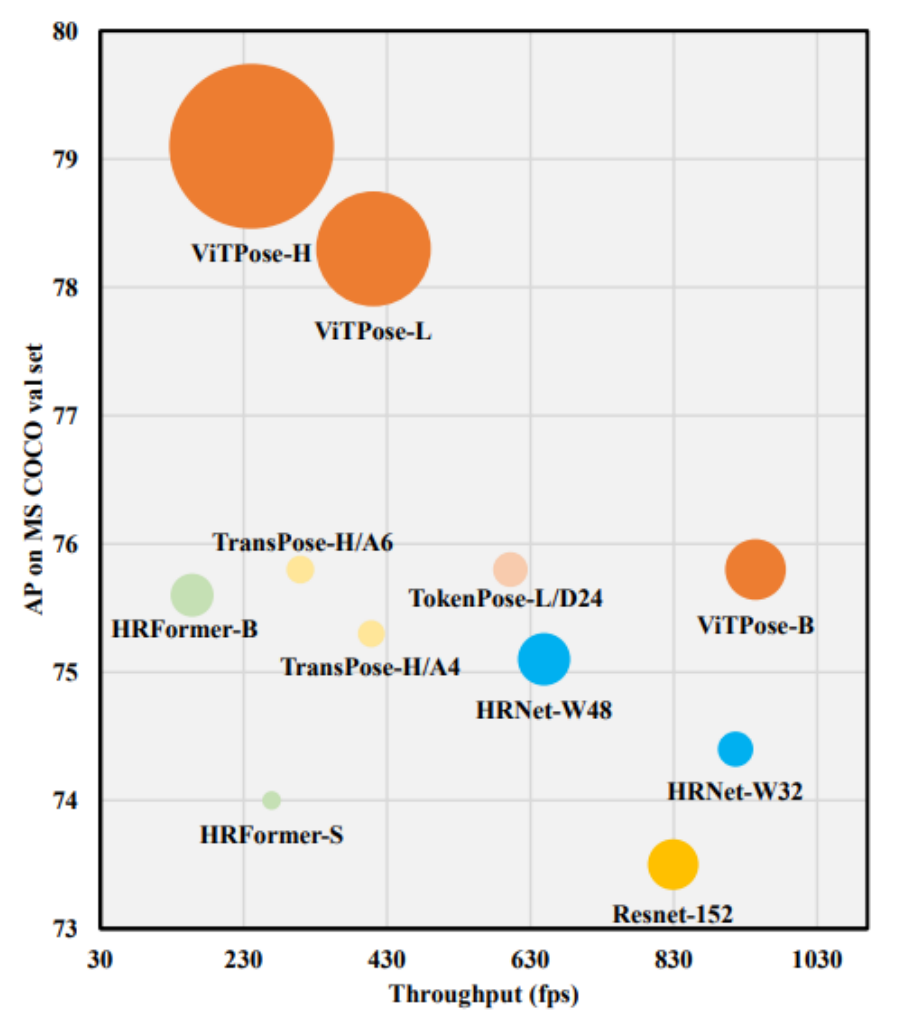
연구 진행 과정

포즈 추정 모델 선정 : ViTPose

사진 속 사람의 자세를 분석하기 위해 ViTPose를 포즈 추정 모델로 선정했습니다. ViTPose는 Vision Transformer 기반으로 높은 정확도로 관절 좌표를 추정할 수 있어 본 연구의 요구사항에 적합하다고 판단했습니다.



The framework of ViTPose.



포즈 데이터 수집: SNS, AI - HUB

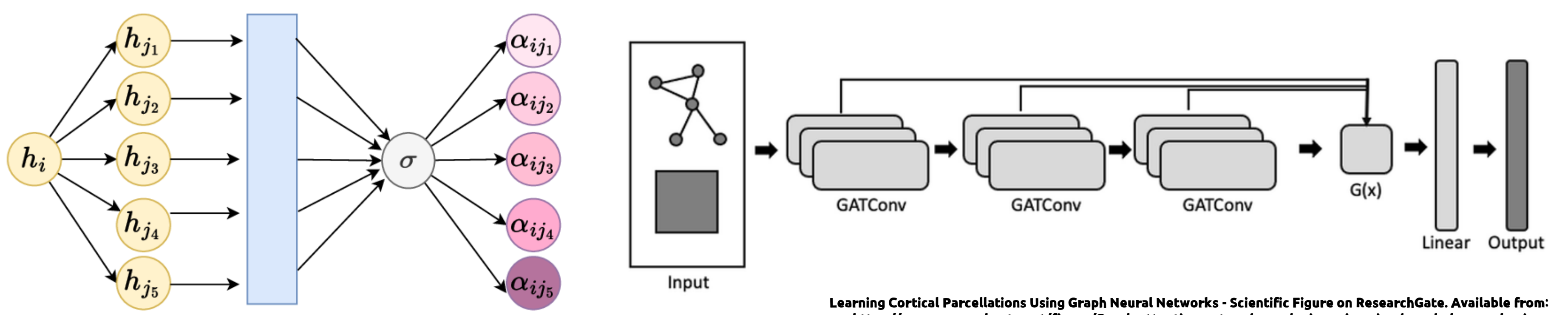
한국인 전신 및 포즈 데이터를 기반으로 하여 수집한 데이터로, 이상적인 자세의 기준을 정의하기 위해 사용되었습니다. 특히, 한국인의 체형과 자연스러운 포즈 특징을 반영한 데이터를 통해, 시스템의 정확성과 현실성을 높이고자 했습니다. 또한 SNS에 공개된 게시물 사진의 좋아요 수, 팔로우 등을 고려하여 수집하였습니다

데이터 영역	영상이미지	데이터 유형	이미지
데이터 형식	jpg	데이터 출처	직접 수집
라벨링 유형	전신 촬영 이미지 데이터 (키폴인트, 세그멘테이션) 고해상도 얼굴영역 이미지 데이터 (세그멘테이션)	라벨링 형식	json
데이터 활용 서비스	한국인 인체 특성을 고려한 다양한 산업군 개발 및 연구 등	데이터 구축년도/데이터 구축량	2023년/45,000개

학습에 사용된 AI hub 한국인 전신 및 포즈 데이터 정보

포즈 학습 모델 결정 : GAT (Graph Attention Networks)

포즈를 학습시키기 위해 딥러닝 모델을 검토한 결과, CNN은 연산 비용이 높고 keypoint 기반 데이터 처리에 비효율적이었기 때문에 제외하였습니다. 반면, GNN은 그래프 형태의 관절 데이터를 효과적으로 처리할 수 있어 적합하다고 판단했습니다. 최종적으로, 관절별 중요도를 다루고 인접 관절 간 관계를 학습하기 위한 GAT(Graph Attention Networks)를 채택했습니다. GAT는 단위 벡터로 각도를 표현하여 연결 정보를 학습할 수 있도록 설계했습니다.



Distributed Node Classification with Graph Attention Networks - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Graphical-representation-of-a-GATConv-layer_fig2_367457060 [accessed 26 Nov 2024]

Learning Cortical Parcellations Using Graph Neural Networks - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Graph-attention-network-employing-a-jumping-knowledge-mechanism-The-network-takes-as_fig2_357321038 [accessed 27 Nov 2024]

결과 및 분석

포즈 추정

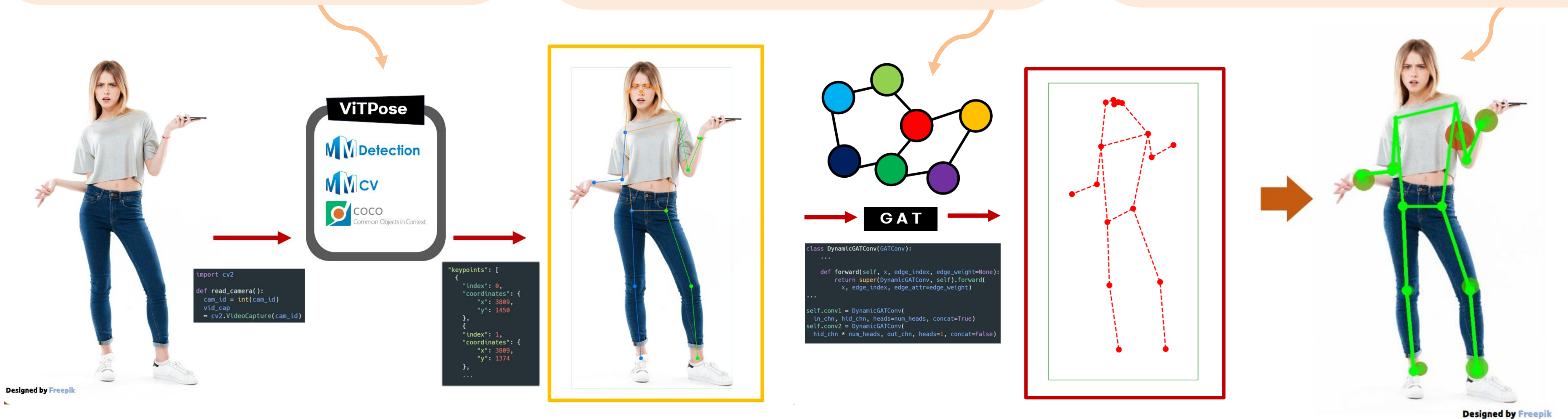
ViTPose, MMPose, MMDetection, OpenCV, COCO 데이터셋을 활용하여 영상 속 사람의 포즈를 추정합니다. 이를 통해 관절별 keypoint 좌표를 추출하고, 추정 결과를 스켈레톤 형태로 시각화합니다.

포즈 평가

추정된 포즈 데이터는 GAT모델로 전달됩니다. 각 관절을 노드로, 관절 간 관계를 엣지로 표현하는 그래프 구조를 생성하며, 인접 관절 간의 방향성을 단위 벡터로 계산하여 모델의 입력 특징으로 사용합니다.

포즈 피드백

GAT 모델은 입력받은 데이터를 기반으로 추정된 포즈를 평가합니다. 평가 결과는 이상적인 포즈와 비교되며, 오버레이 시각화를 통해 사용자에게 직관적으로 피드백을 제공합니다.



오픈소스 URL

<https://github.com/shear99/Programming>

