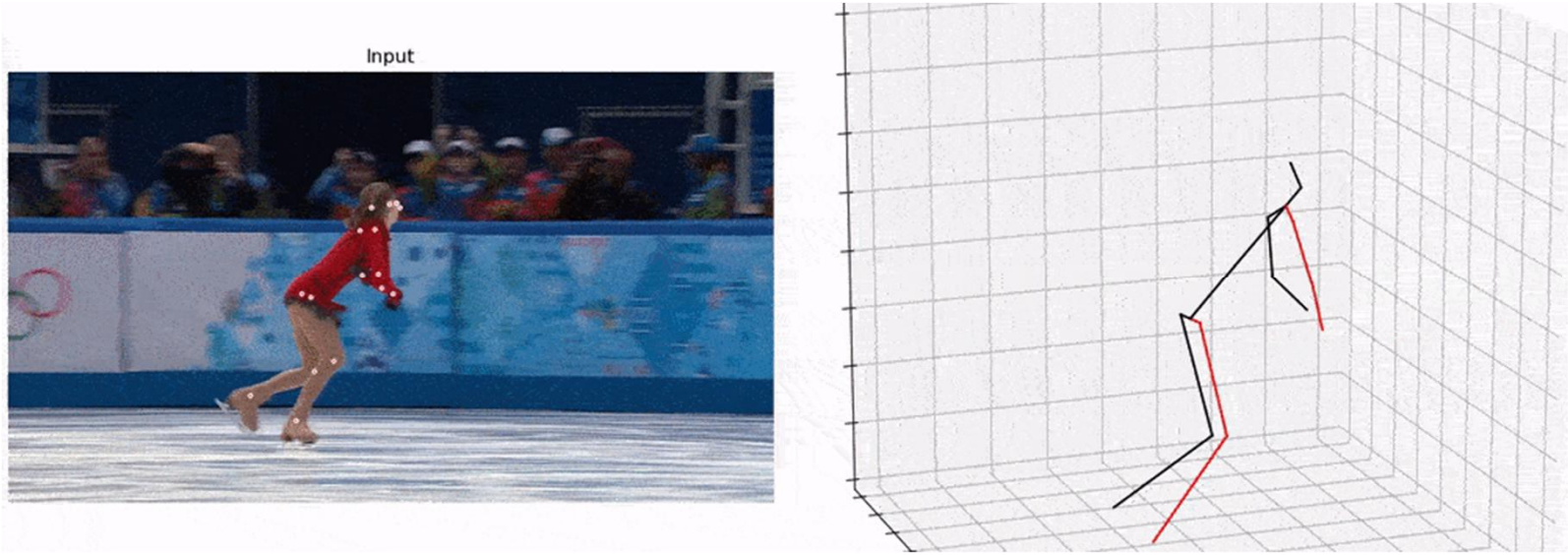
3D Human Pose Estimation with Spatial and Temporal Transformers

Proceedings of the IEEE International Conference on Computer Vision (ICCV 2021)

Human Pose Estimation

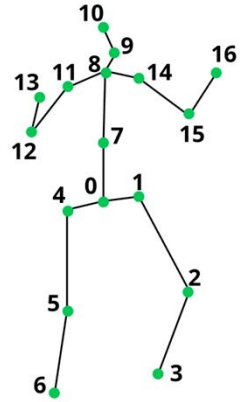
Human Pose Estimation



Basic Concepts of HPE

3D KEYPOINTS AND THEIR SPECIFICATION

0 — Bottom torso
1 — Left hip
2 — Left knee
3 — Left foot
4 — Right hip
5 — Right knee
6 — Right foot
7 — Center torso
8 — Upper torso



9 — Neck base
10 — Center head
11 — Right shoulder
12 — Right elbow
13 — Right hand
14 — Left shoulder
15 — Left elbow
16 — Left hand

mobidev

3 Keywords for HPE

Target : Multi-person VS Single-person

Approach : Top-down VS Bottom-up

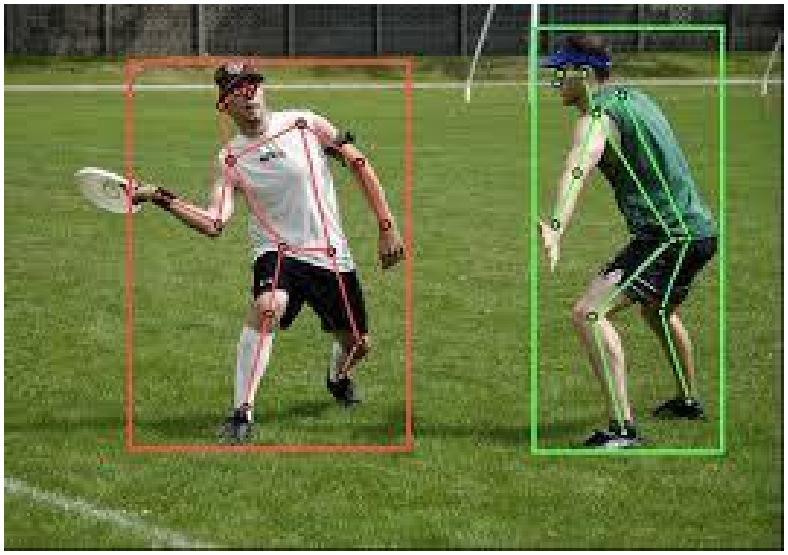
Goal : 2D pose VS 3D pose

Multi-person VS Single-person

Multi-person VS Single-person

Top-down VS Bottom-up

Top-down VS Bottom-up



2D Pose VS 3D Pose

2D Pose VS 3D Pose

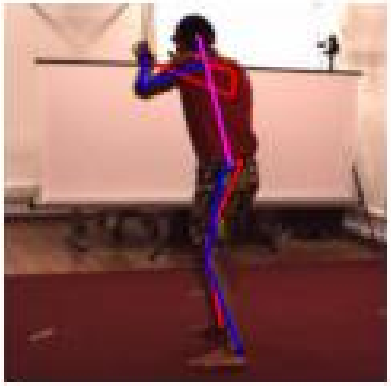
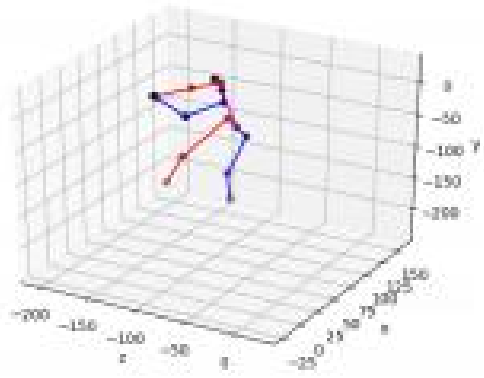
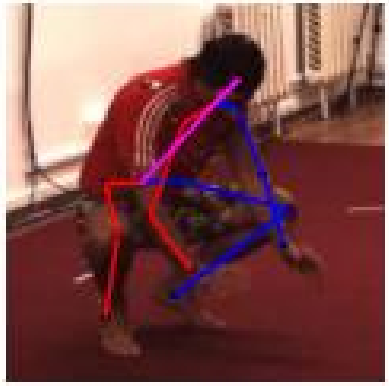
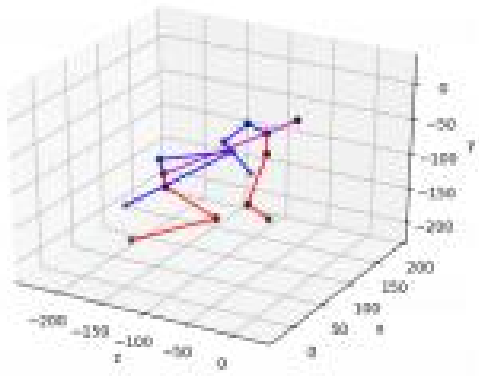
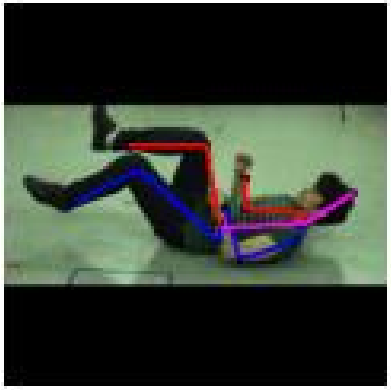
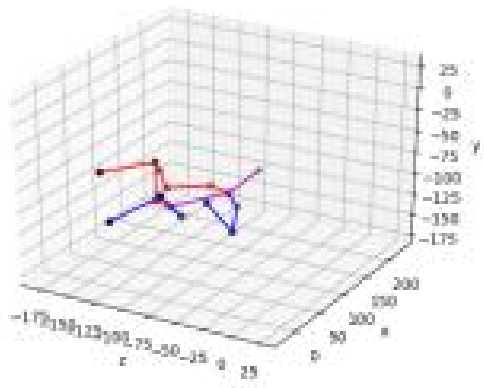
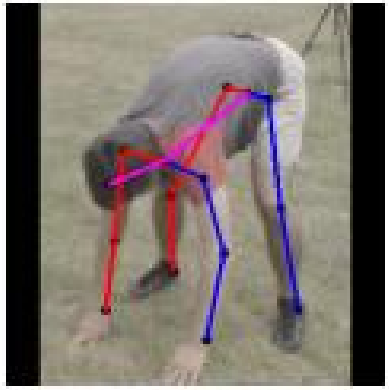
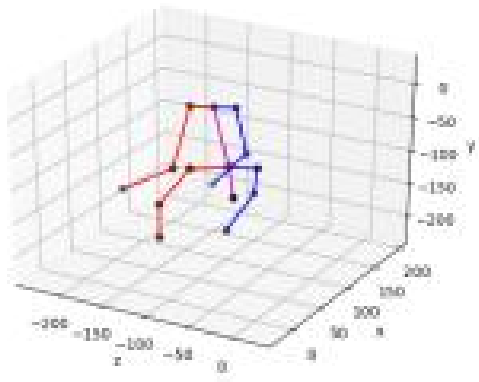
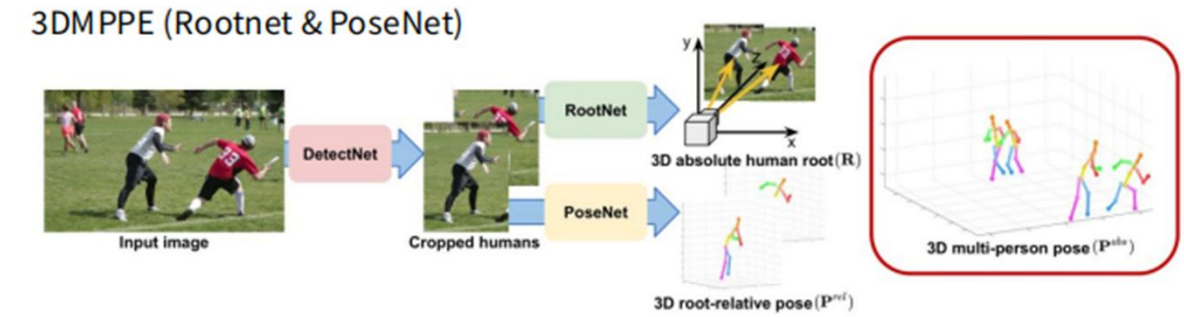
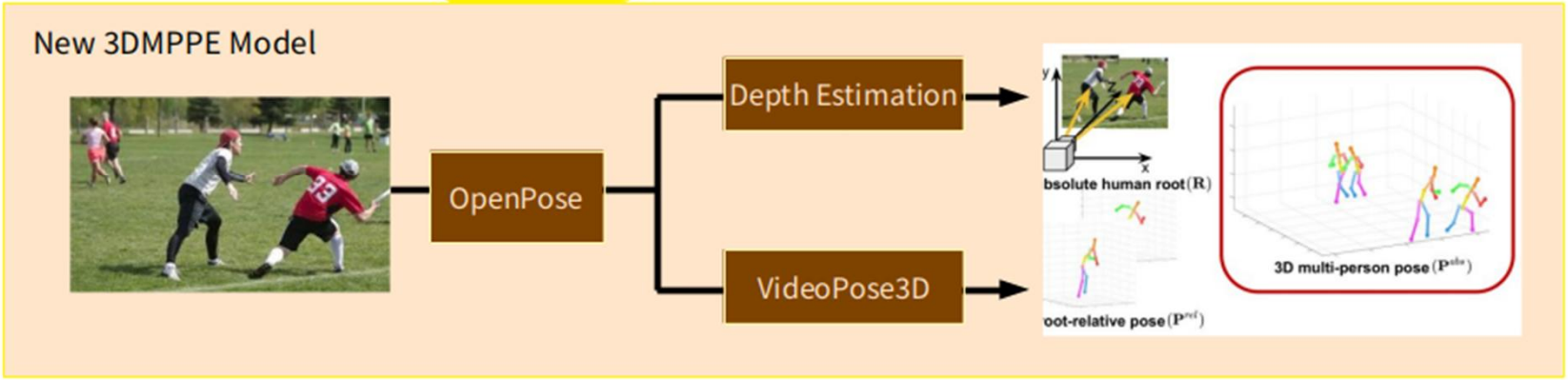









Image 3D Pose



기존 3DMPPE모델에서 DetectNet부분만
Mask-RCNN에서 OpenPose로 변화 시도.
PoseNet부분이 2D-keypoint 정보를 사용하지 않고,
RootNet이 Bounding box를 사용한다는 점에서
OpenPose와 다소 부적합



위와 같은 모델을 제안 할 수 있다. 그러나, 다소 수정 사항이 많아 작은 부분부터 구현 필요하다.

Video 3D Pose