MediaPipe 포즈 감지(Pose)

**MediaPipe**

MediaPipe란 구글에서 제공하는 AI 프레임워크로서, 비디오형식 데이터를 이용한 다양한 비전 AI 기능을 파이프라인 형태로 손쉽게 사용할 수 있도록 제공된다. AI 모델개발 및 수많은 데이터셋을 이용한 학습도 마친 상태로 제공되므로 간편하게 호출하여 사용하기만 하면 된다.

기본적인 얼굴인식 이외에도 Pose 인식 등 다양한 비전AI 기능들이 제공되는데 본 서비스에서는 Pose만 사용한다.

**Pose Land Mark**

포즈 랜드마크의 리스트이다.. 각 랜드마크는 다음과 같이 구성된다.

x와 y : 랜드마크의 좌표는 각각 이미지의 너비와 높이로 [0.0, 1.0]로 표준화된다.

z : 엉덩이 중간 지점의 깊이를 원점으로 하여 랜드마크의 깊이를 나타내며, 랜드마크가 카메라에 가까울수록 값이 작아진다. z의 크기는 x와 거의 비슷한 척도를 사용한다.

Pose Land Mark를 통해 영상 속 사람의 Skeleton Model을 추출할 수 있다. 추출된 Skelton Model의 Pose Land Mark의 반복적인 움직임을 통해 올바른 동작 자세를 추출할 수 있다.

**Pose Classification**

MediaPipe를 활용하면 포즈 분류가 가능하다. 각 운동의 최종 상태( 팔굽혀 펴기를 예로 들면 위, 아래의 위치)에 대한 수백개의 샘플을 추출해낸다.

Pose Land Mark를 특징 벡터로 변환하기 위해 손목과 어깨, 두 손목 사이의 거리와 같이 미리 정의된 포즈 관절 목록 사이의 쌍별 거리를 사용한다. 알고리즘은 거리에 의존하기 때문에 모든 포즈는 변환 전에 동일한 몸통 크기와 수직 몸통 방향을 갖도록 정규화된다.

**Repetition Counting**

Pose Classification를 통해 동작을 분류하고, Repetition Counting을 통해 반복적인 움직임을 관측한다. 그 후 Pose Land Mark을 통해 x,y,z,의 움직임을 수집한다.

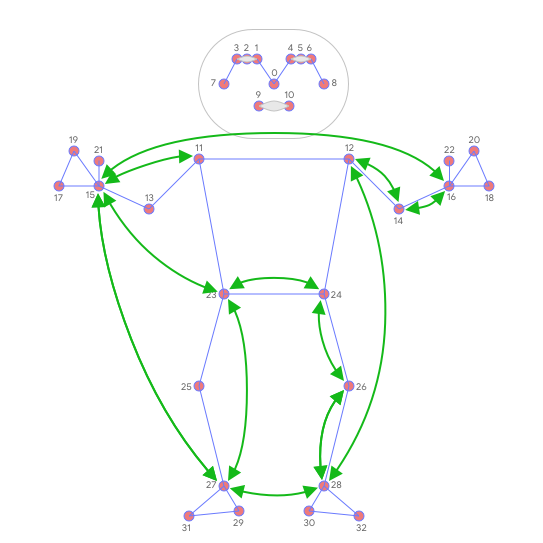
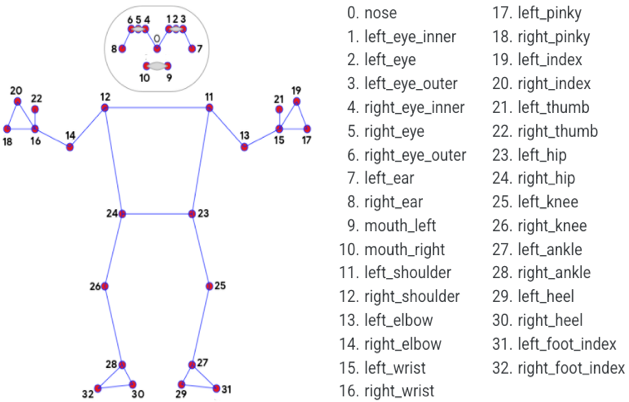


Image 1: Pose Land Mark

Image 2 : Pose Classification