

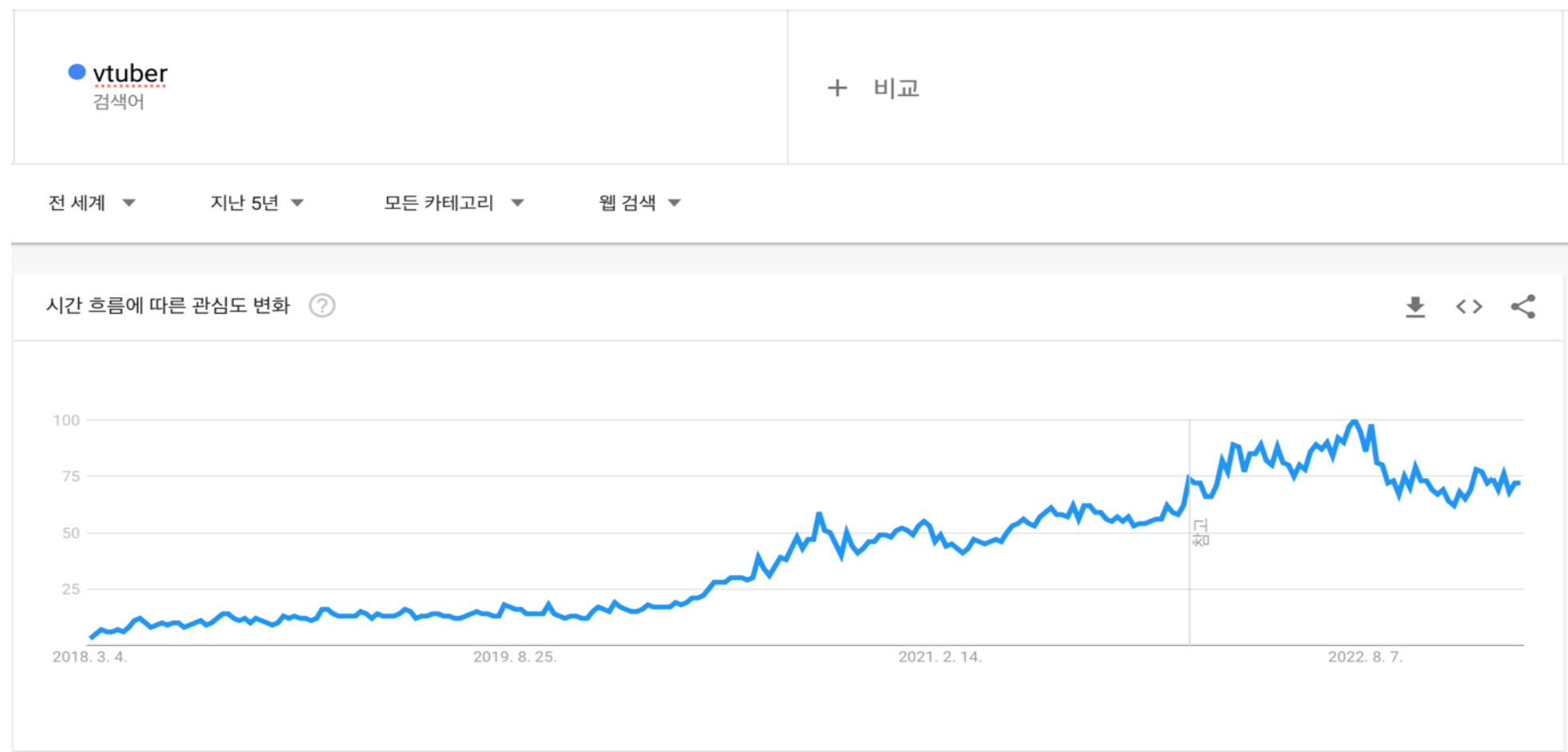
컴퓨터공학과 다인원 포즈 렌더링 프로그램

팀명 : 포즈 마스터즈

팀원 : 유규빈, 함상진

지도교수 : 이계식 교수

목적 및 필요성



인터넷 방송과 메타버스의 기술 발전이 이루어지면서 VR 방송 콘텐츠라는 게 생겨나게 됐고, 그렇게 사람의 행동이나 표정을 대신 표현해주는 캐릭터가 방송을 진행하는 버추얼 유튜버가 등장했다. 버추얼 유튜버라는 게 다소 생소할 수 있는 장르일 수 있지만 구글 트렌드 분석에 따르면 코로나19가 본격적으로 확산한 2020년 상반기부터 버추얼 유튜버에 대한 관심도가 빠르게 증가하는 추세를 보이는 걸 알 수 있고, 전세계 유튜버 일간 슈퍼챗 순위를 버추얼 유튜버들이 상위권을 차지하는 등, VR 방송 콘텐츠에 대한 관심이 뜨겁다는 걸 알 수 있다.

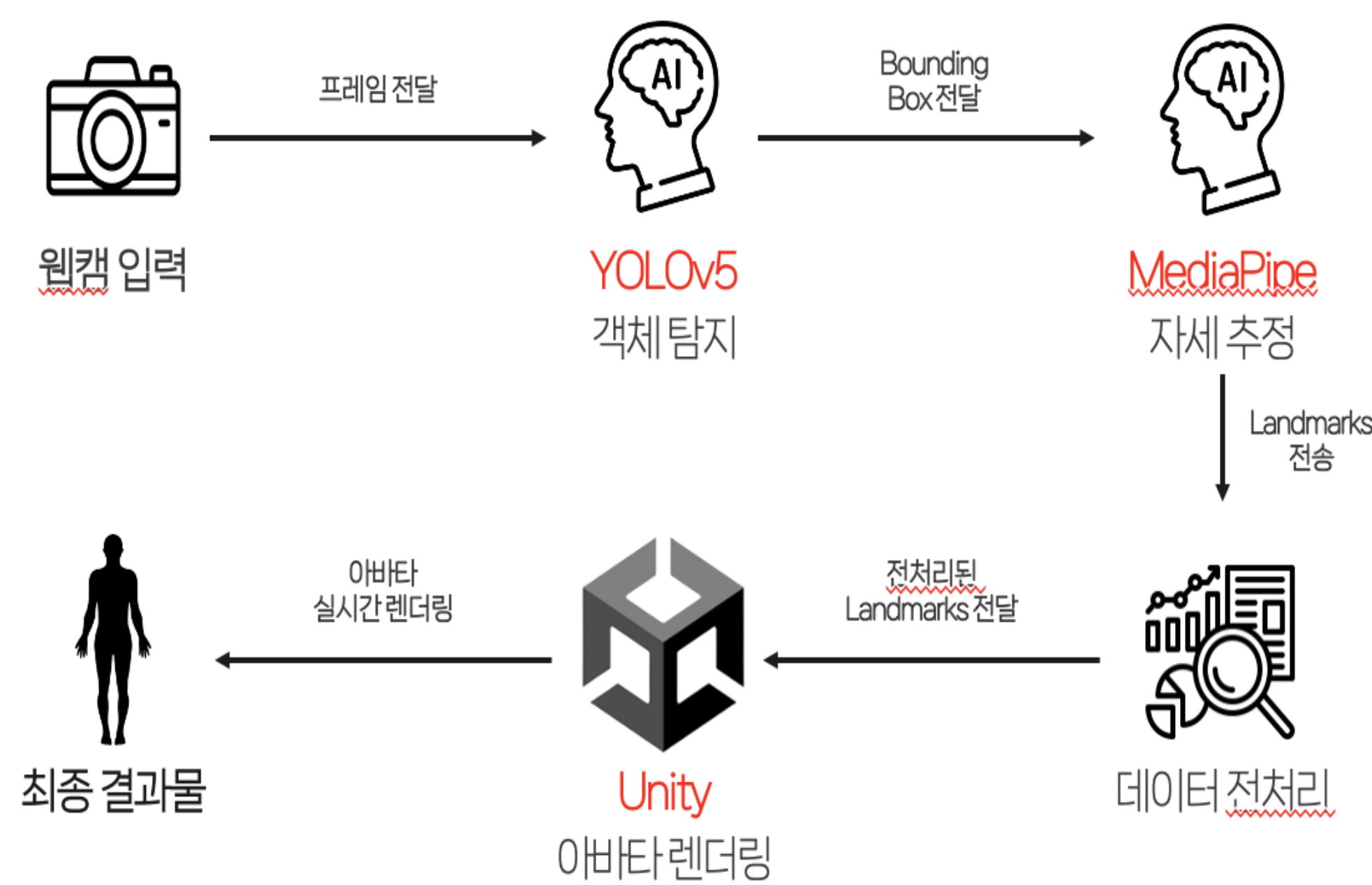
하지만 VR 방송 콘텐츠에서 사용하는 기존 프로그램 같은 경우 표정과 머리 움직임을 제외하고는 인식이 어렵고 한 명밖에 인식이 안 된다는 단점이 있었으며, 프로그램이 아닌 장비를 사용해 머리와 손 움직임 정도만 인식하려 해도 50만원 가량의 가격이 필요해 부담이 컸다.

그래서 기존 프로그램과 장비 사용의 단점을 보완해 장비를 착용하지 않고도 두 명 이상의 전신을 인식하는 게 프로그램의 목적이다.

작품 사진



과제 해결방안 및 수행과정



YOLOv5와 MediaPipe 모두 실시간에서 처리 가능한 가벼운 모델이지만 두 모델을 같이 사용하며 인원 수만큼의 반복 실행으로 인해 속도가 느려졌다.

이를 해결하기 위해 아래 방법들을 사용했다.

- 제한적 객체 탐지 : 카메라 속 사람의 위치가 크게 움직이지 않는다는 사실을 이용하여 매 프레임 사용하던 YOLOv5를 3프레임에 한 번으로 사용을 줄였다.

- 멀티 스레딩: YOLOv5에서 탐지한 인원 수만큼 반복하여 MediaPipe를 사용해야 한다. 이때, 시간비용이 많이 들어 멀티스레딩으로 불필요한 시간을 줄였다.

기대효과

사람의 전신을 인식하기 위해 사용했던 고가의 장비 착용을 대체하여 금전적인 부담이 높은 VR 장비 사용의 진입장벽을 해소할 수 있다. 또한 한 카메라에 1명만 인식하던 기존 프로그램의 제한을 늘려 두 명 이상의 사람 인식이 가능해지고 이에 따라 더욱 다양한 엔터테인먼트 콘텐츠에 활용 가능하다.

실시간으로 자세를 추정하기 때문에 다양한 스포츠에서 운동자세를 분석하는데에 사용할 수 있다. 또한 우리가 사용하는 Holistic 모델 내부에는 Hands 모델도 포함되어 손동작을 인식해 실시간 수화 표현에도 활용할 수 있다.

만약 기존에 존재하는 아바타만을 프로그램에서 활용하는 것이 아닌 원하는 아바타를 생성해주는 기능도 구현한다면 더욱 경쟁력 있는 프로그램이 될 수 있을 것이다.