(1페이지)

안녕하세요 당신의 밸런스를 맞춰드립니다. 의 밸런스가 핵심이조의 발표를 하게 된 “이지평”입니다. 저희 팀원 이렇게 구성되어있습니다. 발표 시작하겠습니다.

(2페이지)

저희는 서론, 모델 구현, 그리고 목적을 이루기 위해서 생각한 방향에 대한 제안과 추가적인 연구방향를 제시하면서 마무리하도록 하겠습니다.

(3페이지)

저는 개인적으로 헬스장을 주기적으로 다니면서 정말 재밌게 운동을 하고 있는 1인입니다

운동을 하면서 많은 분들이 본인의 운동의 밸런스가 맞지 않을 때 종종 다치거나 다음날 담이 오는 등으로 고생을 하곤 합니다. 그렇다고 전문적으로 헬스장 PT를 결제하자니 서울 평균 1회에 7~8만원이라는 높은 비용이 소모됩니다.

PT비용을 들이지 않고 부상을 예방하고 본인의 운동 밸런스를 맞춰주는 프로그램이 있다면

많은 수요가 있을 것이라고 생각하게 되었습니다.

(4페이지)

운동의 좌우밸런스를 맞추기 위해서 2가지의 목적을 설정했습니다.

첫번째 로는 어느정도 자세를 숙지하는 사람들을 대상으로 운동의 좌우 밸런스를 측정하고 그 자세가 현재 적정범위에 들어가는지 아닌지 확인할 수 있어야한다.

두번째는 위 와 같은 과정을 본인이 혼자라도 사용가능하게 하여 혼자라도 자세를 교정하고 부상을 예방하여 고액의 pt 비용없이도 밸런스가 잡힌 운동을 하는 것을 목표로 합니다.

(5페이지)

과제를 수행하기에 앞서서 기존연구 사례들을 조사해 보았습니다. <인공지능 기반의 스마트 헬스케어 운동관리를 위한 어플리케이션 구현> 에서는

인간의 운동동작들을 Posenet을 통해서 POSE ESTIMATION을 진행합니다. 그리고 운동의 강도가 얼마나 강했는지, 그리고 몇회를 진행했는지를 측정해주는 구조입니다.

그리고 상대적으로 가벼운 PoseNet모델을 활용했기에 모바일 기기로도 이식이 가능한 구조를 가진다는 특징을 발견했습니다.

하지만 Posenet은 기본적으로 바텀업 방식이기 때문에 속도측면에서 실시간으로 포즈 에스티메이션이 가능하다는 장점이 있지만 정확도 측면에서 부족함을 보여줬습니다.

(6페이지)

<AI hub 피트니스 자세 이미지 데이터 구축활용 가이드라인>에서는

운동종목별 상세 운동상태를 정의하고, Input 비디오 데이터와 Keypoint 데이터를 동시에 사용하여 모델링을 하고 output으로 운동의 종목과 운동수행정도를 측정해줍니다.

이 연구에서는 같은 데이터를 사용하지만 저희의 목적과는 다른 target값을 가지고 있기에 데이터를 사용한 방식을 차용하되 저희가 필요한 부분과 필요하지 않는 부분을 나누어 참고했습니다.

(7페이지)

기존연구1 에서 바텀업 방식의 장점과 단점이 저희 주제와 상반된다 생각했기 때문에 바텀업의 반대임 탑다운 방식을 채용하여 시간은 좀더 소요되더라도 정확도를 향상시키는 방향으로 결정했고 기존연구 2에서는 저희의 목적에 맞는 target 값을 리라벨링을 해야하는 고민을 해보았지만

그것은 시간적으로도 비용적으로 한계가 있다고 생각하여, 키포인트를 기반으로 좌우밸런스 여부를 판단하는 커스텀 펑션을 만들어 좌우밸런스를 판독하는 것으로 경정했습니다.

(8페이지)

문제를 정의 해보겠습니다. 저희의 목표는 쭉 말씀드렸다시피 다양한 운동동작들에서 좌우밸런스가 잘 맞는 동작과 맞지 않는 동작을 구분해야합니다.

그러기 위해서 정면이 아닌 비스듬한 각도에서 찍힌 영상의 각도를 조정해야하고

좌우 밸런스가 맞는다는 것에 대한 정의가 필요합니다.

(9페이지) 데이터셋

저희가 학습에 사용할 데이터 셋입니다. AI Hub의 피트니스 자세 이미지 데이터로, 총 30종의 동작을 다양한 형태의 피사체를 통해 담은 이미지 데이터입니다. 데이터의 양은 총 20만 클립이며, 5개의 멀티 뷰로 촬영했습니다. 앞서 기존연구2 부분에서의 목적을 이루기 위해 input 데이터는 원본 데이터와 키포인트가 찍혀 있는 데이터가 있습니다. 저희는 모델링을 통해 키포인트를 찍어내는 것이 목표이므로, 키포인트가 찍혀 있는 데이터는 활용하지 않고, 오른쪽 이미지와 같은 원본 이미지 데이터만 학습에 사용할 예정입니다.

(10페이지) 전처리

전처리입니다. 우선 객체를 화면중간에 위치하도록 처리합니다. 여기서 주의해야 할 점은 ‘발이 보이도록’ 처리해야 한다는 점입니다. 이 점은 추후에 저희가 제안하는 방법론에서 굉장히 중요한 부분을 차지합니다. Data augmentation 파트입니다. 배경제거를 통해 객체만 보이도록 처리하고, random flip을 이용해 여러 각도로 flip 시켜줍니다.

(11페이지) 학습1

저희가 학습에 사용할 모델입니다. Backbone 모델로 ‘HRNet’(High resolution network)을 사용할 것입니다. 이는 Top-down 방식의 모델 중 하나로, pose estimation 관련 여러 분야에서 SOTA는 아니지만, 상위권에 위치하고 있는 모델입니다. 기존의 방법과 동일하게 down sampling과 up sampling을 반복하지만, 이를 ‘병렬’적으로 진행함으로써 고해상도와 저해상도의 feature map 간 상호작용이 지속적으로 일어날 수 있게 합니다. Loss function으로는 keypoint를 얼마나 비슷하게 찍었는가를 보는 ‘OKS’(Object Keypoint Similarity)를 사용합니다.

(12페이지) 학습2

성능을 높이기 위한 방법으로 DEKR(Disentangled Keypoint Regression)을 사용할 것입니다. 이는 pixel-wise spatial transformer를 통해 adaptive convolutions을 채택하여 키포인트 영역에 있는 픽셀을 활성화시킨 다음 활성화된 픽셀에서 representation을 학습하게 됩니다. 학습된 representation은 키포인트 영역에 포커스를 맞출 수 있습니다. 쉽게 설명하면, 서로 다른 keypoint를 회귀하기 위한 representation을 분리하여 각 representation이 해당 keypoint 영역에 집중하도록 하는 방법입니다. 이를 통해 각 키포인트별로 집중해서 학습시킬 수 있습니다.

(13페이지)

영상을 찍을 때 지면과 수평이 되지 못하고 비스듬한 각도에서 찍히는 경우가 있을 것입니다. 그럴 때 기준이 되는 축을 정하여 비교를 해야합니다. 이를 위해 필요한 후처리들이 있습니다.

먼저 밸런스 확인을 위해서 운동별로 기준축이 되는 신체부위가 발인지 손인지 구분하고 비교해야 할 주요 신체부위를 지정해놓습니다.

이렇게 한 다음, 운동별로 지정한 신체부위들의 키포인트 좌표값을 이용해 기울기를 구해야합니다. 이 때 기준이 되는 신체부위의 직선은 기준 축으로 지정합니다.

그런 다음 기울기를 비교하여 수평여부 확인 및 각도를 조정하는 순서로 후처리를 진행해야 합니다.

(14페이지)

발이 기준이 되는 운동과 손이 기준이 되는 운동을 분리하고 운동별로 주요 신체부위를 dictionary형태로 지정해놓은 예시입니다.

(15페이지)

모델링을 통해 얻은 키포인트와 운동별로 지정한 주요 신체부위를 통해 좌우 밸런스를 어떻게 정의할 것인지에 대한 custom-Function을 생성하였습니다.

진행과정은 다음과 같습니다.

먼저 운동종목별 밸런스 계산에 필요한 신체부위의 좌표값을 output에서 뽑아냅니다.

그 다음 좌표값을 이용하여 각 신체부위를 이은 직선을 만들어 기울기를 구합니다.

기준축이 되는 직선의 기울기와 주요신체부위 직선의 각도를 파악하기 위해 cosin삼각함수를 이용한 내적식을 응용하여 각도값을 추출합니다.

이러한 각도추출 과정을 모든 신체부위에 대해 병렬적으로 진행한 다음 밸런스 정도와 각도값을 함께 표시해주는 과정입니다.

(16페이지)

이는 기준축과 나머지 축으로 기울어진 방향과 기울어진 정도를 비교해 밸런스 정도를 지정한 예시입니다.

(17페이지)

후처리 과정을 예시를 들어 다시한번 설명드리자면,

사진과 같이 주요 신체부위의 키포인트만을 뽑아내어 좌표값을 받고 이 좌표값을 이용하여 직선을 이은 다음 기울기를 구하게 됩니다.

(18페이지)

그 다음 기준축의 벡터값과 특정 신체부위의 벡터값을 내적한 후 이를 기준선의 길이 곱하기 특정 신체부위 직선의 길이로 나누어주면 cos𝛉 값이 나오게 됩니다. 이때 cos𝛉의 값의 부호를 통해 어느 방향으로 기울어졌는지 파악할 수 있습니다. 예를 들어 cos𝛉가 -1/2이 나왔다면 이는 음의 방향으로 기울어진 것입니다.

다음으로는 cos𝛉를 통해 각도 𝛉를 추출해야합니다. 이때 𝛉가 90도 이상인 경우 시각적으로 어느정도 기울어졌는지 보기 힘들것입니다. 그렇기 때문에 180도-𝛉를 하여 표시해줍니다. 예를 들어 𝛉가 120도가 나왔을 경우 120도 기울어졌다! 보다 60도 기울어졌다!로 표시해주는 것입니다.

이렇게 나온 기울어진 방향과 기울어진 정도를 비교하여 밸런스 정도를 표시하게 됩니다. 예시를 보면 발과 손은 기울기가 다르지만 발과 어깨는 기울기가 같고, 차이가 오차범위 이내이기 때문에 각도와 함께 Good!을 표시해주는 것입니다.

(19페이지) 파이프라인

전체적인 파이프라인은 이렇습니다. input으로 이미지가 들어가면, 앞서 언급했던 DEKR을 통해 각키포인트가 될 수 있는 후보별로 히트맵을 생성합니다. 이것이 HRNet을 통과하면 keypoint를 detection한 좌표를 뽑아내게 됩니다. 이 좌표를 저희가 custom한 Balance function에 동작의 이름과 함께 넣어주면, 연결점들을 이은 각 선들의 각도와 밸런스 정도를 뽑아냅니다. 이것이 저희의 최종 output이 되는 것입니다.

(20페이지) 기여점 및 추가적 연구방향 제시

이 연구의 기여점으로는 사람의 운동이미지를 input으로 사용해서 그 운동이 좌우밸런스가 맞게 행하고 있는 것인지 파악 가능한 파이프라인을 생성했다는 것에 의의를 두고 있습니다. 하지만 아직 부족한 점이 많아 추가적 연구를 진행한다면, 각도조정 후처리 관련 양발이나 양손을 수평의 기준으로 사용하지 못하는 경우에 대해서 좀 더 고민해볼 필요가 있습니다. 그리고 현재는 양발의 직선과 양 무릎의 직선을 보는 형태이지만, 추후에는 발과 무릎을 이은 직선을 생성하여 기준선과 직교하는지 파악하는 기능 또한 추가해볼 예정입니다. 마지막으로 좌우밸런스 여부뿐만 아니라 운동자세의 정확도까지 파악하여 결과를 뽑아내는 기능을 추가하는 방안에 대해서도 생각해볼 것입니다.

(21페이지) 참고자료

저희의 참고자료는 이렇습니다.

이상으로 밸런스가 핵심이조의 최종발표를 마치겠습니다. 감사합니다.