

심층신경망 기반 재실자 이미지 학습을 통한 인체 관절 인식 모델 개발

Human Joints Estimation model through Analysis of Occupant Image based on Deep Neural Network

한지훈¹, 최은지², 문진우², 유용석^{3*}

¹ Department of Computer & Information Communications Engineering, Hongik University
² School of Architecture and Building Science, Chung-Ang University, South Korea
^{3*} Department of Electronic Engineering, University of Incheon (Corresponding author)

Abstract

The aim of this study is to develop a human joint recognition model to provide a basis for measuring the Metabolic rate (MET) of the occupant which is influencing indoor thermal comfort. The Convolutional Neural Network (CNN) as a type of Deep Neural Network (DNN) was used for the development of human joint recognition model. The images of indoor activities and 14 major joint coordinates were collected for model training. The residual block was used for the CNN model and the Percentage of Correct Part (PCP) was used for the performance evaluation. The PCP represents the recognition accuracy of each coordinate and the average PCP of the joint recognition model is 0.75. Based on this accuracy, the model shows a possibility of recognizing human joints from the image.

KEYWORDS

- Thermal Comfort
- MET
- PMV
- Deep Neural Network

1. 서론

1.1 연구의 배경 및 목적

재실자의 생활수준이 향상됨에 따라 실내 환경에서 질병 등 공기 오염에 대한 고려와 함께 쾌적성에 대한 욕구가 증가하고 있다. 이때 실내 열환경(Thermal Quality)은 실내환경의 쾌적도를 결정하는 주요한 요소이다. 열환경은 6가지 물리적 요인에 의해 쾌적도가 결정되며 재실자의 삶의 질 향상과 함께 생산성, 건강 등에 영향을 준다. 이때 고려되는 물리적 요소는 환경적 요소인 온도, 습도, 기류속도, 평균복사온도(Mean Radiant Temperature, MRT)와 개인적 요소인 활동량(Metabolic rate, MET)과 착의량(Clothing, CLO)으로 분류된다. 이러한 환경 및 개인적 요소를 통합적으로 고려하여 실내 쾌적을 나타내는 지표로 평균온열감(Predictive Mean Vote, PMV)을 주로 사용한다. PMV의 주요 요소에 대한 정확한 측정이 가능하면 재실자의 열적 쾌적을 만족하는 실내환경 제어가 가능하다. 하지만 재실자의 개인적 요소는 정확한 측정이 어렵기 때문에 심박동기, 적외선 센서 등 기기를 통한 간접적 측정 방법을 사용하였다. 그로 인해 MET 측정 시 오차가 발생할 가능성이 있어 보다 실질적인 측정 방법의 개발이 필요하다. 이는 사람의 지능적 판단을 대신해 재실자의 MET를 구분할 수 있는 지능형 모델의 개발을 통해 측정이 가능할 것으로 사료된다. 따라서 본 연구는 재실자의 이미지를 통해 MET를 측정하고자 이미지의 인체 주요 관절을 인식하는 지능형 모델을 개발하고자 한다. 개발된 관절 인식 모델은 추후 관절 좌표를 통해 포즈와 MET를 산출하는 모델과 결합하여 PMV 제어에 적용 가능한 MET를 산출할 예정이다.

1.2 연구의 방법

본 연구는 재실자의 이미지를 학습하여 이미지 상의 인체 주요 관절을 인식하고자 한다. 연구의 방법은 다음과 같다. (Fig. 1) 먼저 학습에 사용될 재실자의 실내활동 이미지 500장을 구축하였다. 구축된 이미지 학습을 위해 심층신경망 알고리즘 중 CNN(Convolutional Neural Network)을 사용하여 개발하였다. 마지막으로 개발된 알고리즘의 학습과 발생 오차를 최적화 하는 과정을 진행하였고 최종적으로 모델의 성능을 확인하였다.

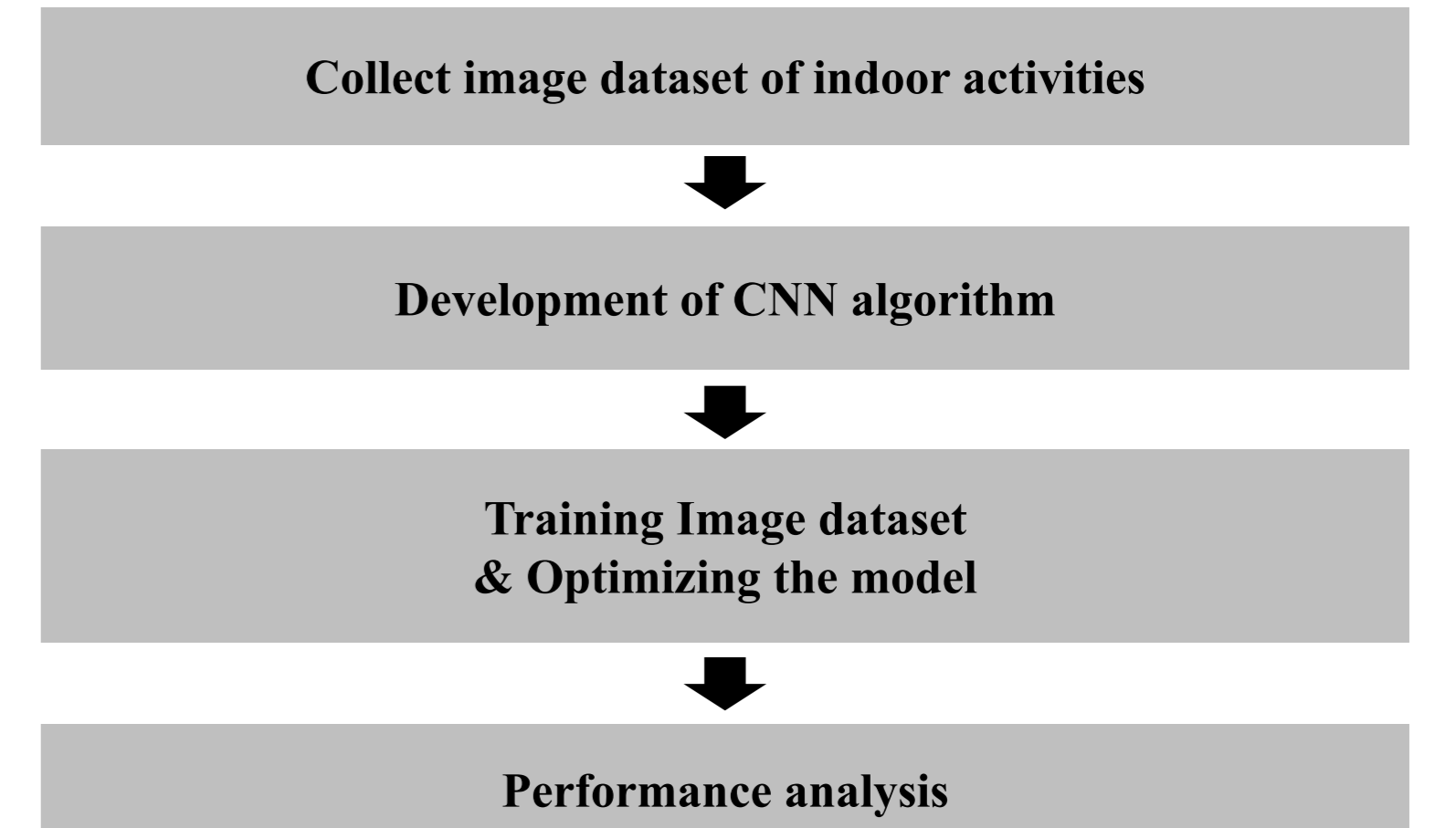


Fig. 1 Process of Joint Coordinates Estimating Model Development

2. 관절 인식 모델 개발

2.1 이미지 데이터 세트

재실자의 이미지 데이터 세트는 Table 1과 같이 AHSRAE 55에서 제시하는 활동 중 실내에서 주로 발생하는 일반적인 실내활동 10가지(sleeping, reclining, seated, quiet, standing, relaxed, reading, seated, writing, typing, filing, seated, filing, standing, walking about)를 선정하였다. 선정된 실내활동에 대한 이미지를 인터넷과 실제 실내환경을 구현한 Mock-up 실험실에서 직접 구축하였다. 총 500개의 데이터를 구축하였고 구축된 이미지는 학습을 위해 인체 주요 관절 14개를 설정하여 해당하는 좌표 값을 학습 시 정답으로 부여하였다. (Table 2) 이때, 14개의 좌표는 양쪽 발목, 무릎, 골반, 손목, 팔꿈치, 어깨와 목, 이마로 설정하였고 (x,y)의 절대 좌표로 구성되어 총 28개의 값을 가진다.

Table 1. Selected Indoor Activities

Activity	Met Unit	
resting	sleeping	0.7
	reclining	0.8
	seated, quiet	1.0
	standing, relaxed	1.2
office activities	reading, seated	1.0
	writing	1.0
	typing	1.1
	filing, seated	1.2
	filing, stand	1.4
	walking about	1.7

Table 2. Joint Coordination Location

Image example	Coordination part
Right	Ankle
	Right Knee
	Right Hip
Left	Hip
	Left Knee
	Left Ankle
Right	Wrist
	Right Elbow
	Right Shoulder
Left	Shoulder
	Left Elbow
	Left Wrist
Neck	
Head	

2.2 관절 인식 모델 구조

관절 인식 모델은 이미지를 입력받아 이미지 상에 인체 관절 좌표를 출력하는 기본 구조를 가진다. 이와 같은 학습을 위해 심층신경망 알고리즘 중 CNN 구조를 포함하는 ResNet을 사용하였다. 이때 CNN 알고리즘은 이미지의 정보를 학습하는데 우수한 성능을 보이는 방법으로 이미지에서 특징을 추출하는 필터를 적용하여 학습한다. 이러한 CNN 구조를 포함하는 ResNet에 Residual Block 개념을 추가하여 적용한 방법이다. Residual Block은 ResNet의 필수 요소로 Fig. 2와 같이 구성된다. Residual Block의 핵심은 입력층과 Convolution Layer의 마지막 단계를 연결해 주어 오차 값이 중간에 소실되지 않게 하여 성능을 향상시키는 역할을 한다. 이와 같이 Residual Block 단위로 관절 인식 모델이 개발되었다. 개발된 모델의 구조는 Fig. 3과 같이 총 Residual Block 30개로 이미지는 CNN Layer를 총 52번 통과하게 된다. 이미지를 입력받아 1차적으로 이미지 특징을 추출한 모델은 최종적으로 14개의 각 Joint 좌표를 결과값으로 내보낸다.

2.3 관절 인식 모델 학습

모델의 학습은 구축한 이미지 데이터 세트의 90%를 학습하고 10%를 평가하는데 사용하였다. 모델의 학습 오차는 모델의 결과 좌표와 정답 좌표의 거리를 계산하여 거리의 차이가 최소가 되도록 학습이 진행된다. 오차는 실제 좌표(p)와 모델이 인식한 좌표(q) 사이의 거리를 측정하는 Euclidean Distance 공식을 사용했다.

$$d(q, p) = \sqrt{(q_i - p_i)^2}$$

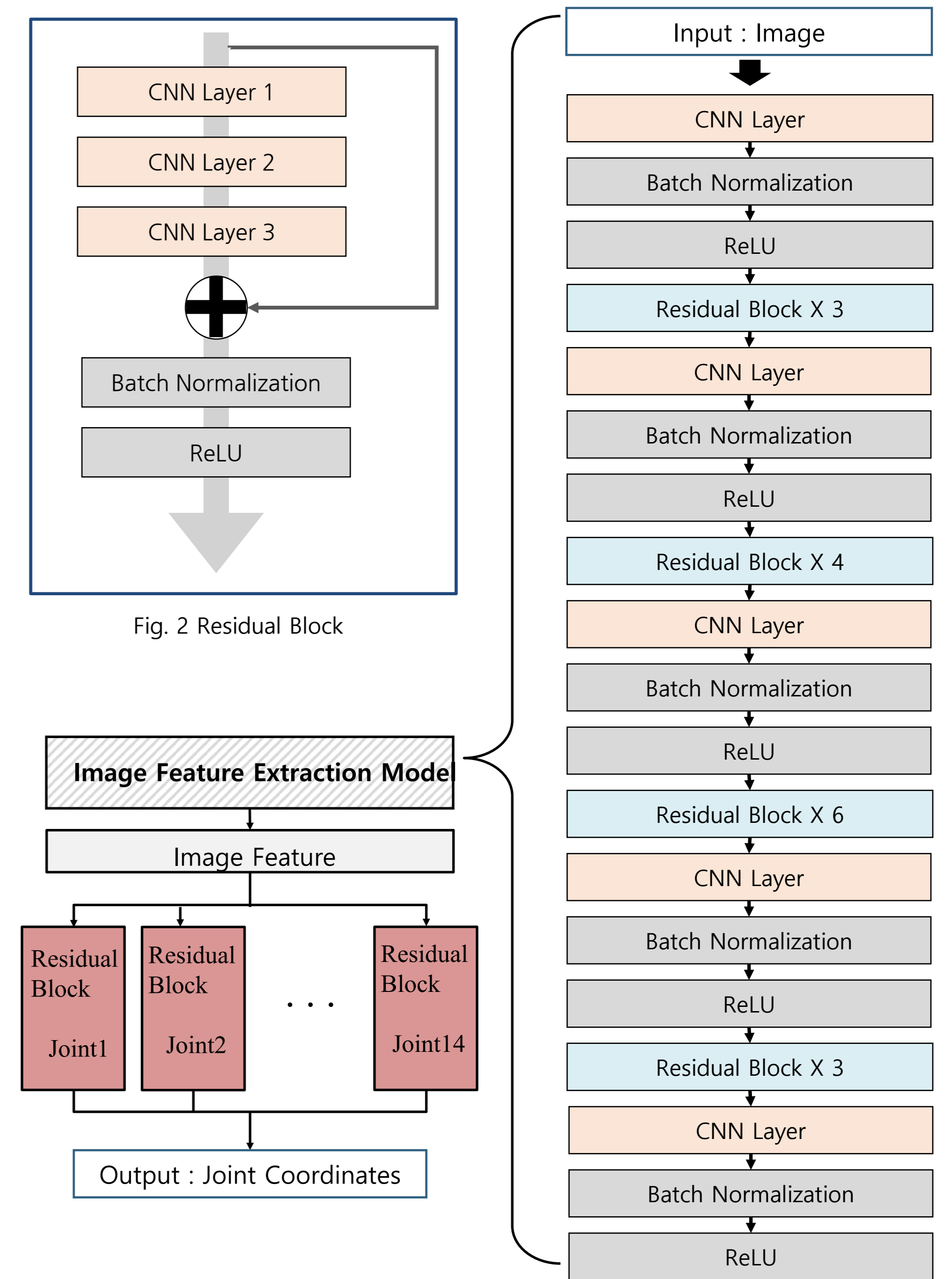


Fig. 3 The Structure of Joint Estimation Model

3. 관절 좌표 인식 모델 학습 및 성능 평가

관절 인식 모델의 성능 평가는 PCP (Percentage of Correct Part) 값의 평균을 사용하였다. PCP는 실제 좌표와 모델의 출력 좌표를 비교하여 정답일 확률을 나타내는 척도로 쓰인다. 이때 실제 좌표를 인접한 좌표와 연결한 Stick 개념을 적용하고 Stick 길이의 정답 영역(correct thresh)을 설정해준다. 정답 영역의 설정 값이 작을수록 모델의 출력 좌표가 정답 범위를 만족하는 것을 강화할 수 있다. 모델이 출력한 좌표에 의해 예측된 Body Part Stick이 형성되고 두 개의 좌표가 모두 정답 영역에 속할 때 올바르게 예측된 것으로 간주한다. 10가지 실내활동 전체에 대한 PCP 평가 결과, 성능은 Fig. 4와 같다. 평균 PCP는 0.75이며 몸통(Torso)의 인식 성능이 0.88로 가장 좋았으며 머리(Head)와 손목-팔꿈치를 연결하는 Lower Arm의 PCP 성능은 약 0.63으로 다소 낮았다. 이러한 결과는 기존 Toshev가 제안한 Deeppose 방법의 결과와 평균 PCP 0.61인 것과 비교하여 관절 인식 성능을 향상시킨 것을 확인하였다.

PCP의 Correct Thresh를 변경하며 성능을 평가한 결과는 Fig.5와 같이 나타났다. 또한 Fig.6과 같이 실내활동 각각의 평균 PCP를 분석한 결과 'sleeping'의 경우 평균 PCP 0.27로 나머지 활동과 비교하여 관절 인식 성능이 저조했다. 또한 대체로 앉아있는 활동('seated, quiet', 'writing', 'typing', 'filing, seated', 'reading, seated')의 경우 서있는 다른 활동에 비해 PCP 값이 0.85 이상으로 더 높게 나타났다. 결과적으로 학습 후 Test 이미지 데이터 세트에 대한 출력 좌표를 나타낸 몇 가지 예시로 각 좌표를 연결한 Stick을 사용해 인식된 좌표를 표현하였다. (Fig. 7)

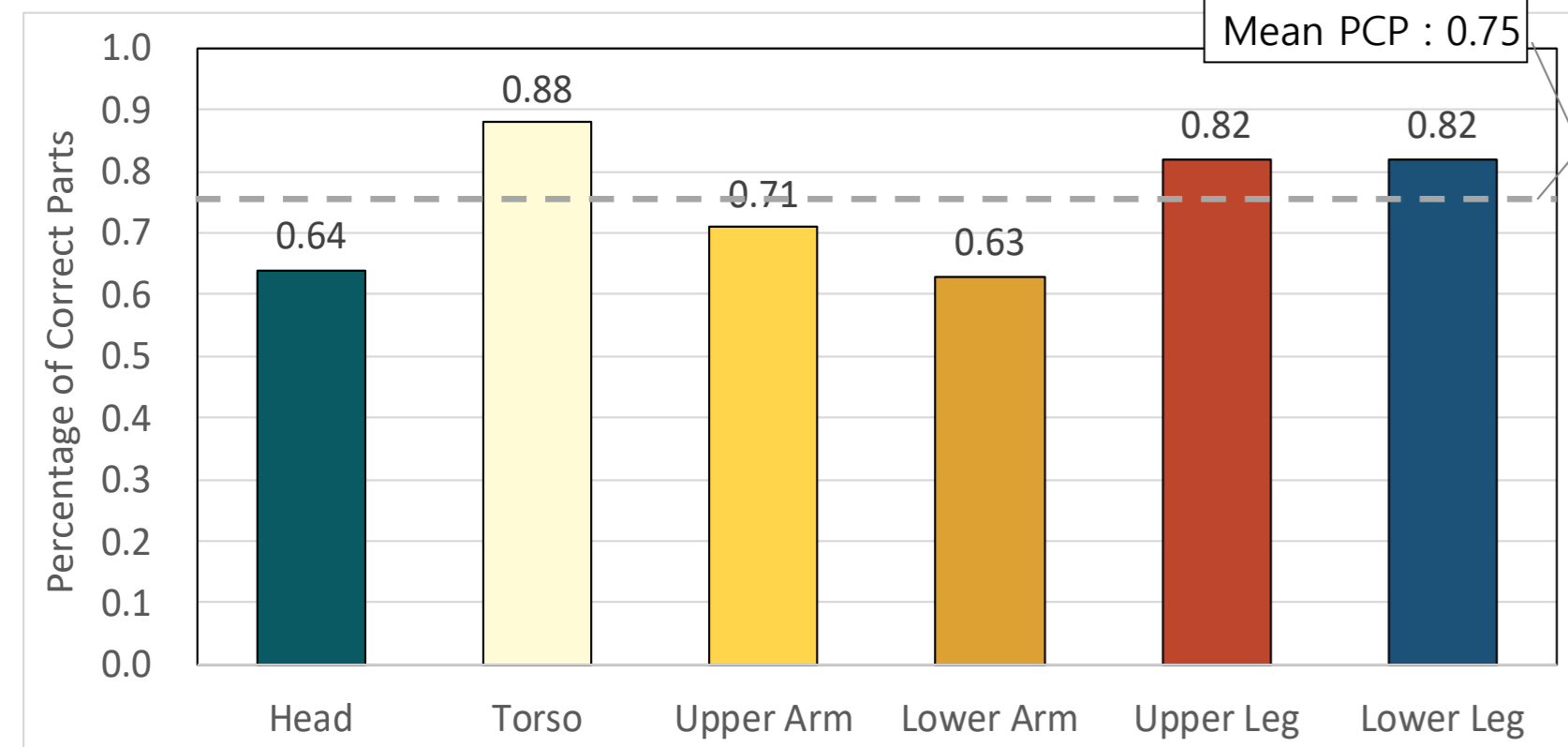


Fig. 4 Mean PCP of each Body Part

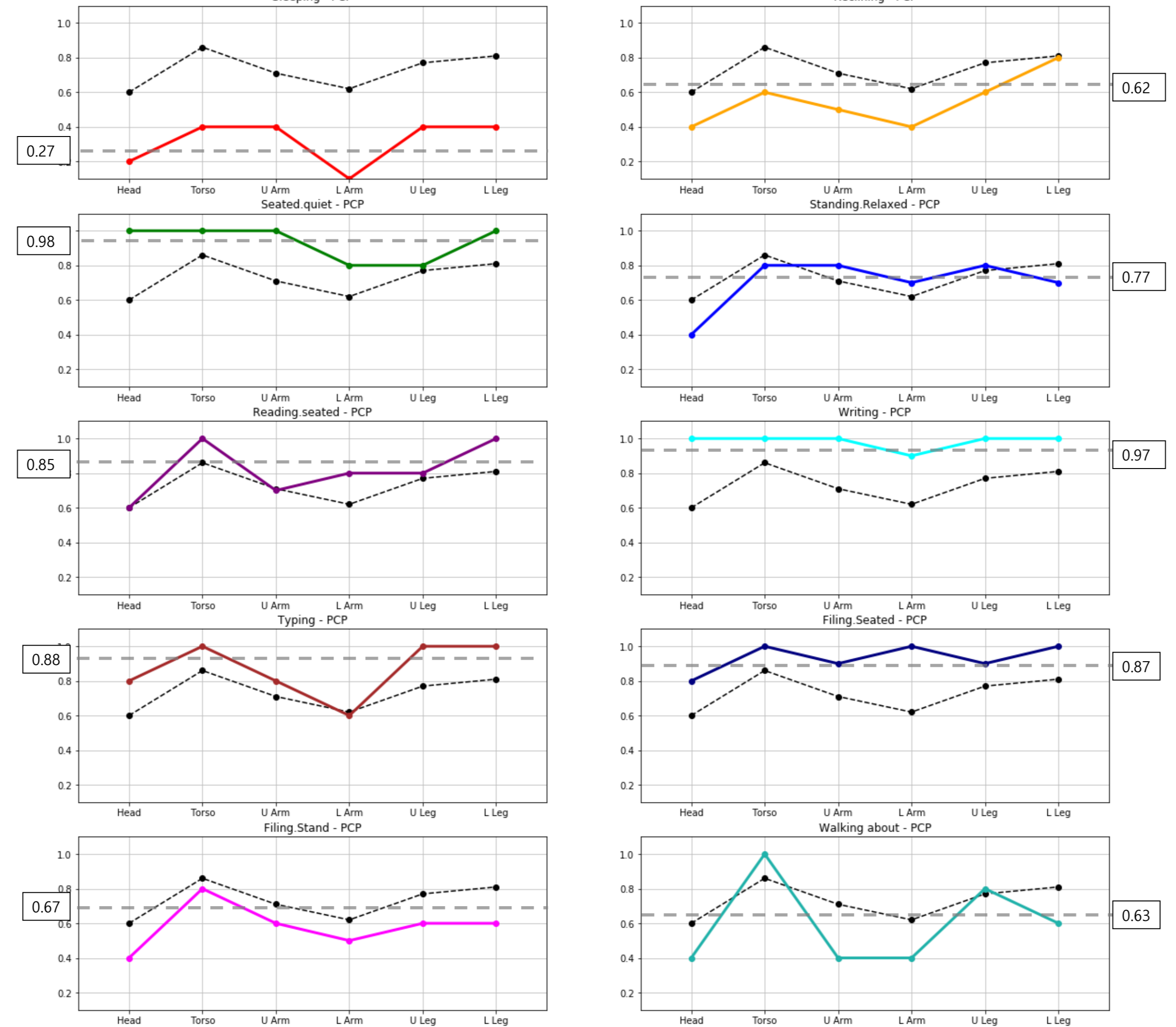


Fig. 6 PCP of each Indoor Activities

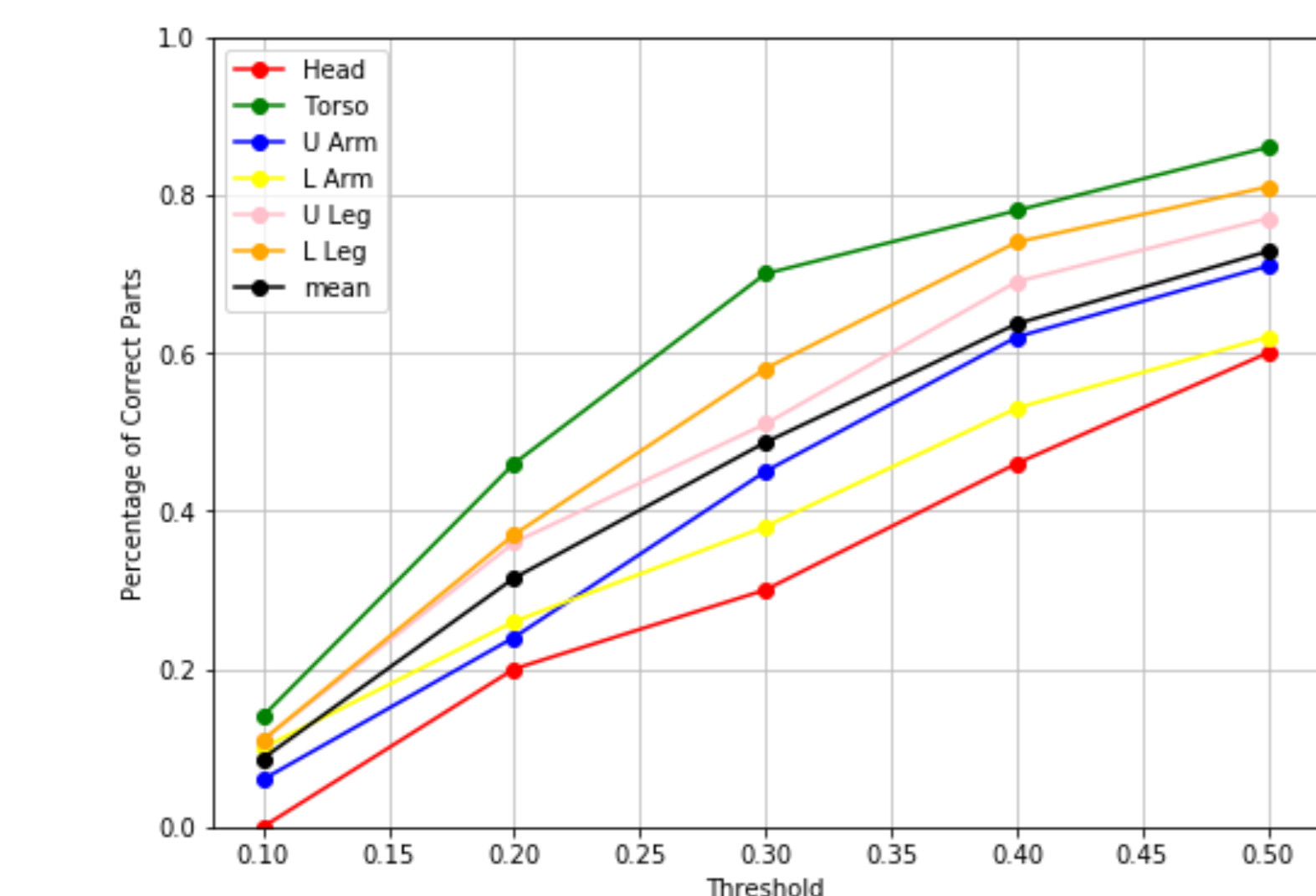


Fig. 5 Mean PCP according to Correct Thresh

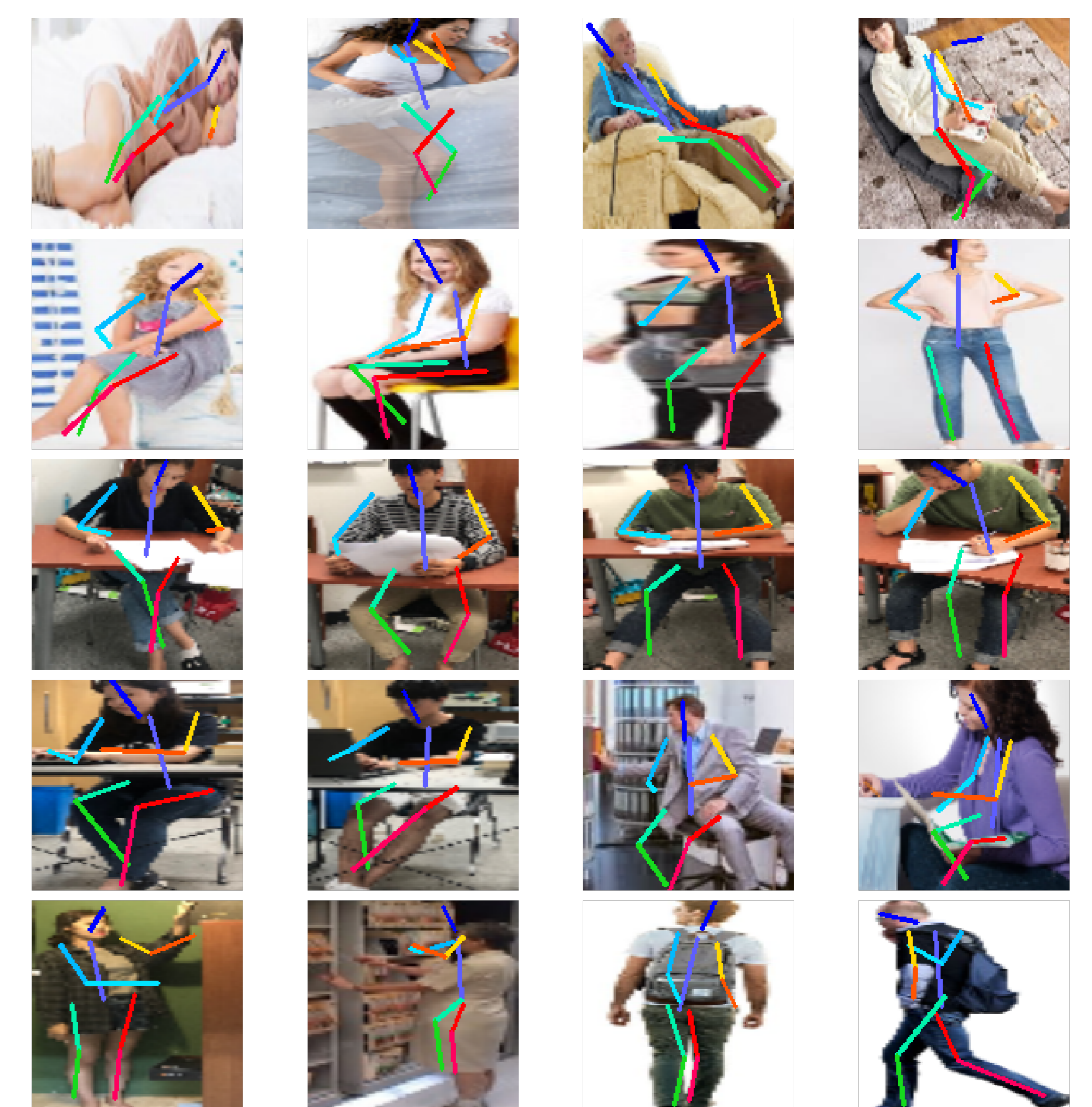


Fig. 7 Examples of Output Coordinates from the Joint Recognition Model

4. 결론

본 연구는 재실자의 실내활동에 대한 MET를 측정하기 위해 재실자 이미지에서 인체 주요 관절을 인식하는 모델을 개발하였다. 모델의 개발을 위해 10가지의 실내활동 이미지 및 관절 좌표 데이터를 구축하였고 CNN 알고리즘을 활용하여 모델을 구성하였다. 학습 결과, 전체 실내활동에 대하여 평균 PCP는 0.75이며 인체 부위 중 몸통 인식 성능이 가장 높았다. 평균 PCP의 성능은 기존 문헌과의 비교를 통해 관절 인식 성능이 향상된 것을 확인하였다. 또한 실내활동 별 성능 평가를 통해 앉아있는 활동이 서있는 활동에 비해 인식이 잘되는 것을 확인하였다.

5. 추후 연구

이러한 결과를 통해 재실자의 이미지를 사용한 관절 인식의 가능성을 확인할 수 있었고 추후 학습 이미지 데이터를 증가시키고 모델의 구조를 변경시키는 등 모델의 성능 향상이 이루어질 예정이다. 또한 개발된 모델의 출력 좌표를 활용하여 재실자의 포즈와 MET를 측정하는데 적용할 수 있을 것으로 기대된다.