

논문 2023-18-06

뇌성마비 환자의 자세 불균형 탐지를 위한 스마트폰 동영상 기반 보행 분석 시스템 (Smartphone-based Gait Analysis System for the Detection of Postural Imbalance in Patients with Cerebral Palsy)

황 윤 호, 이 상 현, 민 유 선*, 이 종 택*

(Yoonho Hwang, Sanghyeon Lee, Yu-Sun Min, Jong Taek Lee)

Abstract : Gait analysis is an important tool in the clinical management of cerebral palsy, allowing for the assessment of condition severity, identification of potential gait abnormalities, planning and evaluation of interventions, and providing a baseline for future comparisons. However, traditional methods of gait analysis are costly and time-consuming, leading to a need for a more convenient and continuous method. This paper proposes a method for analyzing the posture of cerebral palsy patients using only smartphone videos and deep learning models, including a ResNet-based image tilt correction, AlphaPose for human pose estimation, and SmoothNet for temporal smoothing. The indicators employed in medical practice, such as the imbalance angles of shoulder and pelvis and the joint angles of spine-thighs, knees and ankles, were precisely examined. The proposed system surpassed pose estimation alone, reducing the mean absolute error for imbalance angles in frontal videos from 4.196° to 2.971° and for joint angles in sagittal videos from 5.889° to 5.442°.

Keywords : Gait analysis, Postural imbalance detection, Human pose estimation, Tilt correction, Temporal smoothing

1. 서 론

1. 연구 배경 및 목적

보행 분석은 움직임과 자세에 영향을 미치는 신경 발달 장애인 뇌성마비의 임상 관리에 중요한 역할을 하며, 보행 분석 결과는 보행 속도, 보폭, 보속 및 사지 움직임을 포함한 환자의 보행 패턴에 대한 중요한 정보를 제공하고 있다 [1]. 특히 보행 분석은 다음과 같은 측면에서 뇌성마비 임상 관리에 큰 도움이 된다. 첫째, 뇌성마비 장애의 중증도를 정량화하고 시간 경과에 따른 변화를 모니터링함으로써 상태 중증도를 평가한다. 둘째, 특정 보행 편차나 비정상적인 움직임 패턴 등을 감지함으로써 잠재적인 보행 이상을 식별한다. 셋째, 보행 개선과 통증 감소를 목표로 하는 물리 치료, 교정, 수술 등의 개입 계획 및 평가를 위해 보행 분석을 활용한다. 넷째, 환자 상태의 변화를 식별하기 위해 정기적이고 장기적인 보행 분석을 수행하여 향후 비교를 위한 기준을 제공한다.

이처럼 뇌성마비 환자가 꾸준한 재활 치료를 받을 때, 환자의 정확한 치료 상황과 자세의 개선 상태를 파악하기 위해서 가장 정확한 방법은 환자가 병원에 방문하여 측정하는 것이다 [1, 2]. 하지만, 이 방법은 비용과 시간이 많이 들어가기에 환자가 직접 방문하는 경우를 간소화하여 환자의 자세 교정 정도를 꾸준히 편하게 확인하는 시스템의 필요성이 대두되고 있다. 특히 스마트폰을 통한 동영상 데이터 획득의 편의성 증가와 영상 기반으로 사람 관절을 추정하고 분석하는 딥러닝 모델의 발전, 공개로 많은 학습 데이터를 구축하지 않고도 다양한 의료 분야들에 적용되어 좋은 성과를 내고 있다 [3, 4].

본 논문에서는 ResNet 기반의 기울기 판별모델 [5]을 통해 영상의 기울어짐 정도를 판단하여 영상의 수평을 맞추어 영상 분석에 용이하도록 하였고, human pose estimation 모델인 AlphaPose [6]와 human pose estimation 강화 모델인 SmoothNet [7]을 활용하여 jitter를 제거함으로써, 스마트폰 동영상만으로도 재활 환자의 보행 시 자세를 분석할 방법을 제시하고, 그에 따른 자세 불균형 정도를 추정한다. 보행 정면 영상 (frontal)과 보행 측면 영상 (sagittal)에 대해 분석하였으며, 정면 영상에는 척추와 어깨, 척추와 골반 간의 관계를 추정하여 어깨와 골반의 불균형 각도 (imbalance angle)를 추정한다. 측면 영상에서는 척추와 허벅지의 각도, 무릎 각도, 발목 각도를 추정한다. 이를 통해서 의료 실무에서 사용할 수 있는 지표들 (어깨 불균형 각도, 척추와 허벅지의 각도, 무릎 각도, 발목 각도)을 분석한다. 본 논문에서 제

*Corresponding Authors (ssuni119@gmail.com; jongtaeklee@knu.ac.kr)
Received: Feb. 27, 2023, Revised: Mar. 24, 2023, Accepted: Apr. 17, 2023.
Y. Hwang: Kyungpook National University (Undergraduate Student)
S. Lee: Kyungpook National University (MS Student)
Y. S. Min: Kyungpook National University School of Medicine (Assoc. Prof.)

J. T. Lee: Kyungpook National University (Asst. Prof.)
* 본 논문은 정부 (교육부)의 재원으로 한국연구재단 기초연구사업의 지원을 받아 수행된 연구임 (No. NRF-2020R111A3070636).
* 본 논문은 교육부 및 한국연구재단의 4단계 BK21 사업 (경북대학교 컴퓨터학부 지능융합 소프트웨어 교육연구단)으로 지원된 연구임 (419990214394).

시하는 새로운 시스템을 이용하여 기존의 pose estimation만을 이용하는 방식에 비해서 오차율이 정면 영상에서는 4.196°에서 2.971°, 측면 영상에서는 5.889°에서 5.442°로 더 나은 분석 결과를 보여준다.

2. 관련 연구

2.1. Human pose estimation

Human pose estimation은 이미지나 비디오에서 사람의 관절 또는 부위의 위치를 추정하는 것이 목적인 컴퓨터 비전의 주요 과제 중 하나이다. human pose estimation을 통해서 semantic keypoints를 검출하여 각 관절의 위치를 추정하는데, 추정된 값은 행동 인식과 보행 분석에 중요한 역할을 한다 [8]. 환자의 보행 분석에서도 마찬가지로, 관절 위치를 추정할 수 있어야 보행에서 관절의 각도와 관절 간의 관계를 분석할 수 있기에 최신 human pose estimation 모델을 사용하였고, 그중 top-down 방식의 AlphaPose와 bottom-up 방식의 OpenPose를 비교하여 설명한다.

AlphaPose [6]는 top-down 방식을 이용해서 사람을 검출한 뒤에 관절을 추정하는 방식을 사용하였다. 하지만 기존의 top-down 방식과는 다르게 검출기의 의존도를 낮추고 estimation과 tracking을 동시에 학습시켰다. 또한 전체적인 파이프라인을 병렬로 실행이 가능한 5단계의 파이프라인으로 구현하여서 실시간으로 human pose estimation이 가능하게 하였다. OpenPose [9]는 bottom-up 방식을 이용하여 관절을 먼저 추정하고 관절 간의 관계를 파악해서 자세를 추정하는 방식이다. 각 관절의 위치정보의 분포를 파악한 뒤 해당 분포를 통해서 각 신체 부위 간 연관성 정보를 벡터로 파악한다. 그리고 위치정보와 연관성 정보를 통해서 관절이 어떻게 연결되는지 찾아낸다. COCO test-dev 2015 데이터 셋 [10]에 대하여 AP (Average Precision)를 비교하였을 때 OpenPose의 경우 AP@0.5는 84.9%, AP@0.75는 67.5%인 반면, AlphaPose의 경우 AP@0.5는 89.2%, AP@0.75는 79.1%이며, AlphaPose가 더 다양한 방법의 추적기를 제공하기에 본 연구에서는 AlphaPose를 기반으로 시스템을 구현한다.

2.2. 기울기 추정 모델

스마트폰 동영상의 경우 촬영 당시에 촬영자의 실수나 흔들림으로 영상이 기울어질 수 있다. 영상의 기울어짐을 수정하기 위한 연구로, Hong et al. [5]은 주어진 영상에서 기울기를 추정하는 방법을 제안했다. 기울기 추정 모델의 경우 네트워크에 가변 공간적 (adaptive spatial) 풀링 레이어 (pooling layer)를 추가하여 영상의 다중 스케일 특징을 동시에 고려할 수 있게 하여 이미지 전체의 기울기를 추정하는 모델이다. 본 논문에서는 해당 모델을 사용하여 스마트폰 촬영에서 발생하는 기울어짐 문제를 해결한다.

2.3. Temporal Smoothing

대부분의 human pose estimation 방법들은 매 프레임마다 관절별 위치 추정 오류를 최소화하려고 하기에, 관절 추정 위치 값에서 급격한 변화가 자주 발생한다. 이에 따라

분석 결과에 악영향을 끼칠 수 있기에 이를 안정화할 과정을 제안한다. 각 관절 위치정보의 프레임별 상관관계를 파악하여 보정하기 위한 방법으로 temporal smoothing 방식을 이용한다. Temporal smoothing 방식 중 SmoothNet [7]은 최신 제안된 동영상에서의 human pose estimation 강화 네트워크로, human motion video 환경에서 기존 방식들은 output jitter와 같이 프레임 전반에 걸쳐 다양한 추정 오류를 발생시키는데, 거의 보이지 않거나 가려진 동작의 경우 여러 관절의 추정 위치가 연속 프레임 시퀀스에 대한 실제 값에서 크게 벗어나 결과에 영향을 주는 jitter를 발생시키는 점을 해결하기 위해 제안되었다. 이는 시간만을 고려하는 강화 네트워크를 이용하여 모든 관절의 긴 범위의 시간적 관계를 학습하여 전신의 natural smoothness 특성을 모델링한다.

2.4. 영상 기반 환자의 보행 분석

최근 파킨슨병 환자의 보행을 일반 카메라 영상만으로 정량적으로 분석한 연구에서, OpenPose human pose estimation 모델을 기반으로 걸음 분석을 위해 추가로 background에서 점이나 선과 같은 background reference를 추출하여 깊이 추정을 하였다. 그리고 이를 통해서 발의 속도를 계산하여 걸음을 분석하였다 [11]. 해당 선행 연구는 일반 카메라 영상을 이용하여 환자의 걸음걸이를 분석한다는 맥락에 있어서 본 논문과 비슷하지만 몇 가지 차이점을 보인다. 첫 번째 차이점은 영상 촬영 환경에 대한 자유도이다. 선행 연구에서는 정해진 공간에서 고정된 카메라로 정형화된 상황에 대해 촬영한다. 하지만, 본 논문에서는 스마트폰으로 보호자가 임의로 촬영하기에 정해진 규격이 없으며, 촬영되는 공간은 영상마다 다르고, 카메라 또한 고정된 경우가 적다. 두 번째 차이점은 해당 선행 연구는 환자의 걸음 속도에 기반하여 분석하는 반면, 본 논문에서는 영상 내의 관절의 좌표에 기반하여 관절 간의 거리 및 각도를 분석한다는 점이다.

II. 자세 불균형 탐지를 위한 보행 분석 시스템

그림 1과 같이 제안하는 자세 불균형 탐지를 위한 보행 분석 시스템은 이미지 기울기 보정, human pose estimation, temporal smoothing, 보행 자세 불균형 탐지로 구성되어 있다. 영상 기울기 보정 후, AlphaPose를 사용하여 human

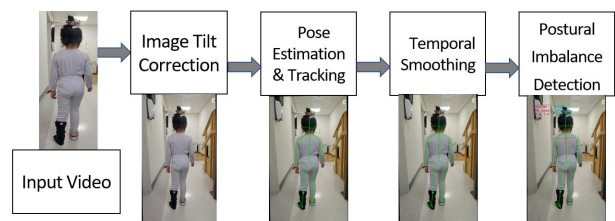


그림 1. 자세 불균형 탐지를 위한 보행 분석 시스템
Fig. 1. Gait analysis system pipeline for postural imbalance detection

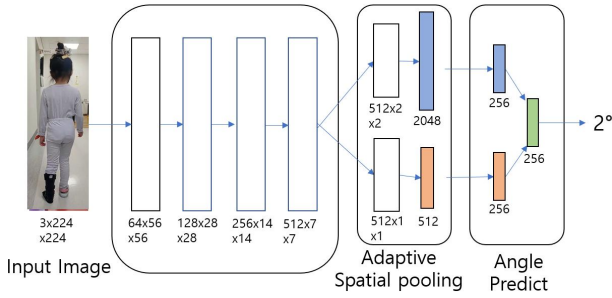


그림 2. 기울기 추정 모델 파이프라인
Fig. 2. Slope estimation model pipeline

pose estimation과 tracking을 진행, 그리고 추정된 semantic keypoints를 SmoothNet을 거쳐 jitter를 제거한 후, 자세 불균형 탐지를 위해 정면 영상은 척추와 어깨, 골반의 불균형 각도를, 측면 영상에서는 척추와 허벅지, 무릎, 발목의 각도를 추산 및 검증한다.

1. 영상 기울기 보정

영상 기울기 보정을 위한 기울기 추정 모델은 ImageNet으로 pretrain한 ResNet32를 backbone으로 이용한다. 기존 ResNet32와 다른 점은 그림 2과 같이 ResNet32에서 하나의 가변 공간적 풀링 레이어를 추가해서, 1x1과 2x2 두 개의 가변 공간적 풀링 레이어로 대체하여 특징 맵을 만든다는 점이다.

각각의 풀링 레이어를 이용하여 구해낸 영상 전체의 기울어짐 정보와 특정 부분의 기울어짐 정보를 결합한 벡터를 통하여 최종적인 각도를 예측한다. 기울기 추정 모델을 학습하기 위한 loss 함수는 L1 loss (absolute difference)를 사용하며, optimizer는 adam, 미니 배치 크기는 32, 학습률은 0.0001, 5e-05, 1e-05로 20 epoch마다 학습률을 낮추면서 학습한다. 이때 네트워크의 후반부 레이어들 (9개의 residual block)의 가중치 파라미터들을 초기화한 후 학습 데이터로 추가 학습한다. 하지만 Hong et al. [5]이 제안한 학습 방법과 달리 MPII 데이터 셋 [12]과 World Cities 데이터 셋 [13]에서 체의 기준을 정하고 직접 선별한 이미지 7,229장을 사용하였으며, 기존의 8장으로 나누어서 학습했던 것과 달리, 학습 데이터를 증강하기 위해 -20~20도 사이의 랜덤한 각도 20장과 원본 1장 총 21장으로 나누서 학습했다.

동영상의 프레임별 각도를 추정하기 위해 매 프레임 단위로 이미지를 추출한 다음, 추출한 이미지를 기울기 추정 모델에 입력하여 영상의 기울어짐 각도 추정값을 출력으로 얻는다. 그리고 추정된 각도의 반대로 영상을 회전하여 영상을 보정한다. 그림 3은 영상 보정 전후 예시 결과를 보여준다.

2. Human pose estimation and tracking

두 번째 단계는 AlphaPose를 활용하여 사람의 관절 위치를 추정하고 추적한다. 그림 4과 같이 입력 영상을 받아 사람을 찾고, 찾은 사람의 바운딩 박스 중에서 기준치 이상으로 겹치는 박스를 찾아 제외한 뒤, pose estimation과 사람



(a) (b)
그림 3. 기울기 보정 영상 예시 : (a) 원본 영상, (b) 기울기 보정 영상
Fig. 3. Example of slope correction image : (a) original image, (b) slope correction image

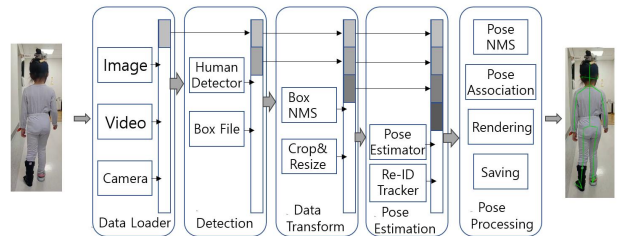
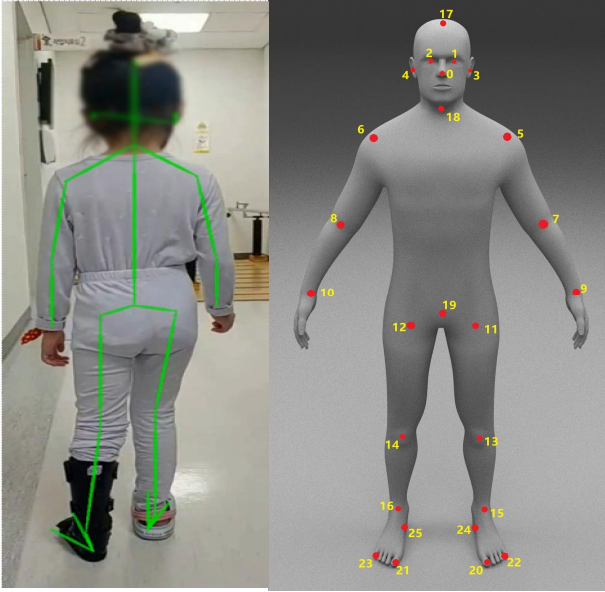


그림 4. Alphapose 파이프라인
Fig. 4. Alphapose pipeline

재인식 기반 추적을 통해 같은 사람을 추적한다. 그리고 찾은 자세 위치 중에서 가장 점수가 높은 자세 위치만 남겨서 해당 자세 위치를 입력 이미지 위에 그려서 출력한다. AlphaPose는 학습 데이터와 정의된 관절의 위치와 수에 따라 COCO dataset [10], Halpe dataset (26 keypoints, 136 keypoints) [6], COCO WholeBody dataset (133 keypoints) [14] 기반으로 학습된 모델을 제공한다. 이 중 그림 5와 같이 26개 keypoint 기반의 Halpe dataset으로 학습된 모델을 본 연구에서 이용하였으며, Halpe26 모델은 사람의 관절을 직관적으로 표현하며 척추를 나타낼 수 있는 관절들이 있어 다른 관절과 척추와의 상관관계를 표현하기 좋다는 장점이 있다. 또한, 여러 사람에 대한 자세 정보 처리와 과검출된 사람 정보를 줄이기 위해, AlphaPose에서 제공하는 추적을 이용하였다. 사람 재인식 기반 추적기, 검출기 기반 사람 추적기, 그리고 PoseFlow기반 사람 추적기가 제공되며, 본 연구에서는 재인식 기반 추적기를 사용하였다.

3. Temporal Smoothing

단순히 AlphaPose만 사용한 semantic keypoints를 정면 영상과 측면 영상 분석에 사용하기에는 key points가 안정적이지 않고 떨림이 컸다. 이런 경우 오류가 누적되어 관절간의 관계를 계산한 결과에 jitter가 생긴다. 그림 6은 정면



(a) (b)

그림 5. AlphaPose 영상 예시 : (a) AlphaPose 출력, (b) Halpe26

Fig. 5. Example of AlphaPose image : (a) output of AlphaPose, (b) Halpe26

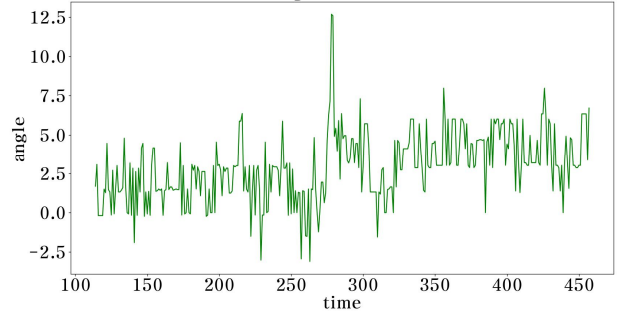
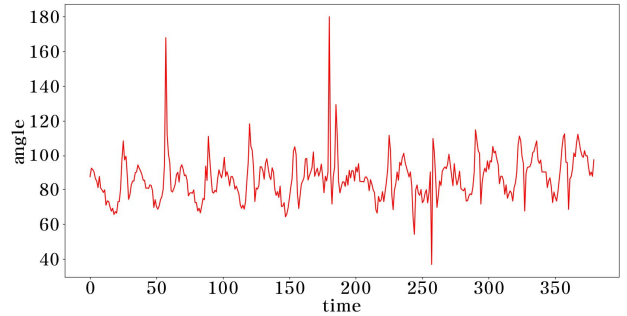
영상의 우측 어깨 불균형 각도를 시간에 따라 나타낸 그래프로 jitter가 크게 발생한 모습을 볼 수 있다. 그림 7의 경우에도 측면 영상의 좌측 발목 각도를 시간에 따라 나타낸 그래프로 jitter가 크게 발생한 모습을 볼 수 있다. SmoothNet이 추정 결과에 바로 적용하여 사용이 가능하고 jitter를 없애는 데 좋은 성능을 가졌기에 사용하였다.

SmoothNet에서 제공되는 2D pretrained model은 Human3.6M [15]를 데이터 셋으로 사용하였다. 본 연구에서는 2D pretrained model 중에서 pose estimator로서 HRNet [16]과 RLE [17] 기반으로 학습된 모델을 사용한다.

4. 보행 자세 불균형 탐지

보행 자세 불균형 탐지를 위해 영상의 종류에 따른 분석 방법이 제시된다. 영상의 종류는 정면 영상과 측면 영상으로 두 가지이며, 정면 영상에서는 어깨, 골반의 기울어진 정도를 분석하며, 측면 영상에서는 허벅지와 척추 각도, 무릎 각도, 발목 각도를 분석한다.

Halpe26을 나타내는 그림 5-(b)에서 영상 분석에 사용될 관절은 정면 영상에서는 5번 (왼쪽 어깨 관절), 6번 (오른쪽 어깨 관절), 11번 (왼쪽 골반 관절), 12번 (오른쪽 어깨 관절), 18번 (목 중심), 19번 (치골)이다. 측면 영상에서는 11번 (왼쪽 골반 관절), 12번 (오른쪽 어깨 관절), 13번 (왼쪽 무릎 관절), 14번 (오른쪽 무릎 관절), 15번 (왼쪽 발목 관절), 16번 (오른쪽 발목 관절), 18번 (목 중심), 19번 (치골), 20번 (왼쪽 엄지발가락), 21번 (오른쪽 엄지발가락), 24번 (왼쪽 발꿈치), 25번 (오른쪽 발꿈치)이다.

그림 6. 정면 영상의 우측 어깨 불균형 각도 추정 그래프
Fig. 6. Right shoulder imbalance angle estimation graph of frontal image그림 7. 측면 영상의 좌측 발목 각도 추정 그래프
Fig. 7. Left ankle angle estimation graph of sagittal image

$$\begin{aligned}
 x_{diff_1}, y_{diff_1} &= (x_{11} - x_{12}, y_{11} - y_{12}) \\
 x_{diff_2}, y_{diff_2} &= (x_{21} - x_{22}, y_{21} - y_{22}) \\
 d &= \det \begin{bmatrix} x_{diff_1} & x_{diff_2} \\ y_{diff_1} & y_{diff_2} \end{bmatrix} \\
 t_x, t_y &= \det \begin{bmatrix} x_{11} & y_{11} \\ x_{12} & y_{12} \end{bmatrix}, \det \begin{bmatrix} x_{21} & y_{21} \\ x_{22} & y_{22} \end{bmatrix} \\
 x_{cross}, x_{cross} &= \det \begin{bmatrix} t_x & t_y \\ x_{diff_1} & x_{diff_2} \end{bmatrix} / d, \det \begin{bmatrix} t_x & t_y \\ y_{diff_1} & y_{diff_2} \end{bmatrix} / d.
 \end{aligned} \tag{1}$$

4.1 정면 영상 분석 자세 불균형 탐지

정면 영상에서는 어깨와 골반의 불균형 각도를 계산하여 어깨와 골반이 기울었는지 분석한다. Halpe26의 관절 모델에서 어깨, 척추, 골반 위치는 다음과 구한다.

- 어깨 : 5번과 6번 관절의 연결 선
- 골반 : 11번과 12번 관절의 연결 선
- 척추 : 18번과 19번 관절의 연결 선

어깨와 골반의 불균형 각도는 기준선 (baseline) 벡터와 각도를 얻고자 하는 관절의 연결 선 벡터, 두 벡터를 통해 계산한다. 이때 기준선은 척추와 어깨, 척추와 골반이 만나는 교점에서 척추와 수직인 직선이다. 어깨의 불균형 각도는 다음과 같이 구한다. 척추 직선은 18번 관절 (x_{11}, y_{11}) 과 19번 관절 (x_{12}, y_{12}) 을 이은 선, 어깨 직선은 5번 관절 (x_{21}, y_{21}) 과 6번 관절 (x_{22}, y_{22}) 을 이은 선이다. 두 직선의 교점은 크라메르 법칙을 이용하여 수식 (1)을 통해 구한다.

이때 $x_{diff}, y_{diff}, t_x, t_y$ 는 크라메르 법칙 (Cramer's rule)을 나타내기 위한 매개변수이다. 크라메르 법칙은 유일한 해를 가지며 변수와 방정식의 수가 같은 연립 일차 방정식의 해를 구하는 공식으로, x_{cross}, y_{cross} 는 각각 교점의 x, y 좌표이다.

교점의 좌표를 구하였을 때 왼쪽 어깨의 불균형 각도를 구하기 위한 벡터 v_j 와 v_h 를 기준선과 교점 (x_{cross}, y_{cross}) , 5번 관절 (x_{21}, y_{21}) 과 수식 (2)를 통해 구한다. 수식 (2)에서 s 는 척추의 기울기, i 는 기준선의 기울기, b 는 기준선의 y 절편, $f(x)$ 는 기준선의 수식, H 는 현재 관절과 같은 x 값을 가진 기준선 위의 점의 좌표, H_x 와 H_y 는 각각 H 의 x, y 좌표, v_h 는 교점과 H 를 이은 벡터, v_j 는 교점과 현재 관절을 이은 벡터를 나타낸다. 이때 왼쪽 어깨와 오른쪽 어깨는 같은 기준선과 연결선을 사용하기에 부호만 다른 같은 크기의 불균형 각도를 가진다. 골반에서의 불균형 각도를 계산하기 위해서는 수식 (1)에 B 를 11번 관절 (x_{21}, y_{21}) 과 12번 관절 (x_{22}, y_{22}) 로 대입하면 골반의 교점을 구할 수 있고 이를 수식 (2)에 대입하여 골반의 불균형 각도를 구하기 위한 v_h 와 v_j 를 구할 수 있다.

$$\begin{aligned} s &= (y_{12} - y_{11}) / (x_{12} - x_{11}), \\ i &= -1/s, \\ b &= y_{cross} - i \times x_{cross}, f(x) = ix + b, \end{aligned} \tag{2}$$

$$\begin{aligned} H &= (x_{21}, f(x_{21})), \\ v_h &= (H_x - x_{cross}, H_y - y_{cross}), \\ v_j &= (x_{21} - x_{cross}, y_{21} - y_{cross}). \end{aligned}$$

이렇게 구한 v_h 와 v_j 를 수식 (3)에 대입하여 불균형 각도 (θ_{imba})를 구할 수 있다.

$$\theta_{imba} = \text{acos}\left(\frac{v_h \cdot v_j}{\|v_h\| \|v_j\|}\right) \times \frac{180}{\pi}. \tag{3}$$

정면 영상 분석에서의 문제가 되는 부분 중 하나는 관절 좌표의 겹침이다. 영상 내에서 환자가 뒤돌아설 때는 그림 8-(b)와 같이 관절들이 겹치거나 너무 근접하는 상황이 발생한다. 그림 8-(a)은 정면 영상 우측 어깨 불균형 각도 추정 그래프에서의 회전 시 오류 영역을 박스로 표시한 것으로 오류가 발생하는 모습을 보인다. 따라서 영상 내에서 환자가 뒤돌아서는 부분은 분석 시에 제외했다. 계산 후 나온 어깨와 골반의 불균형 각도를 그림 9와 같이 분석 영상 왼쪽 상단에 OpenCV를 사용하여 직관적으로 나타냈다. 관절을 나타내는 색과 같은 색으로 표현해주었고, 왼쪽은 어깨의 불균형 각도, 오른쪽은 골반의 불균형 각도를 표현하며, 그 아래로 왼쪽과 오른쪽 중 기울어있는 방향을 알려준다. 파란색과 빨간색 막대는 기울어있는 수치를 가시화한 것이며, 맨 아래 있는 숫자는 기울어있는 수치를 나타낸다.

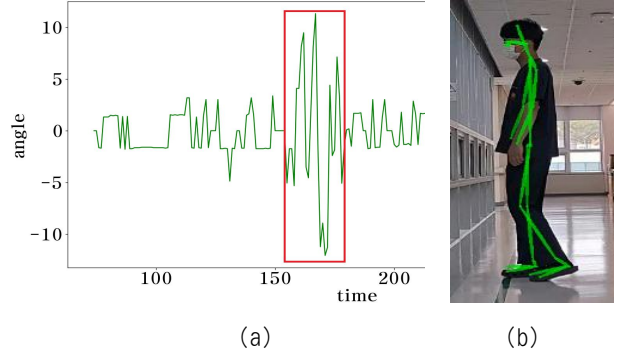


그림 8. 정면 영상 뒤돌아올 때 오류 예시 :
(a) 정면 영상 우측 어깨 불균형 각도 추정 그래프에서의 오류,
(b) 오류 발생 시의 AlphaPose 영상
Fig. 8. Turning error example of frontal image :
(a) Error in right shoulder imbalance angle estimation graph of frontal image, (b) AlphaPose image when error occurs



그림 9. 정면 영상 내 분석 표시
Fig. 9. Analysis display of frontal image

4.2 측면 영상 분석

측면 영상에서는 허벅지와 척추의 각도, 무릎의 각도, 발목의 각도를 계산하여 각 관절의 각도를 분석한다. Halpe26의 관절 모델에서 허벅지, 척추, 무릎, 발목, 발꿈치는 다음과 같다.

- 척추 : 18번과 19번 관절의 연결 선
- 왼쪽 허벅지 : 11번과 13번 관절의 연결 선
오른쪽 허벅지 : 12번과 14번 관절의 연결 선
- 왼쪽 종아리 : 13번과 15번 관절의 연결 선
오른쪽 종아리 : 14번과 16번 관절의 연결 선
- 왼쪽 무릎 : 11번, 13번, 15번 관절의 연결 선
오른쪽 무릎 : 12번, 14번, 16번 관절의 연결 선
- 왼쪽 발목 : 13번과 15번 관절의 연결 선, 20번과 24번 관절의 연결 선
오른쪽 발목 : 14번과 16번 관절의 연결 선, 21번과 25번 관절의 연결 선
- 왼쪽 발꿈치 : 24번 관절 점
오른쪽 발꿈치 : 25번 관절 점

각 관절의 각도는 두 직선 벡터가 이루는 각도이다. 왼쪽 발목 각도는 15번-13번 관절 벡터를 v_1 , 24번-20번 벡터를 v_2 라고 할 때, 수식 (3)에서 v_h, v_j 를 v_1, v_2 로 바꾼 수식 (4)를 이용하여 구한다. 수식 (4)에서 θ_{joint} 는 관절의 각도를 나타낸다.

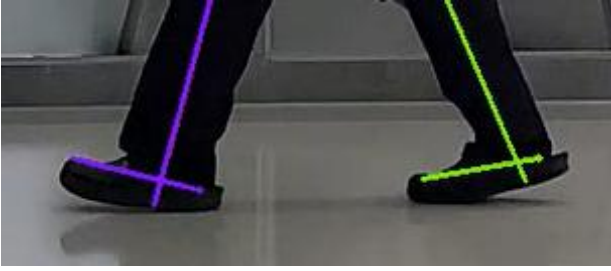


그림 10. 측면 영상 발목
Fig. 10. Ankle of sagittal image

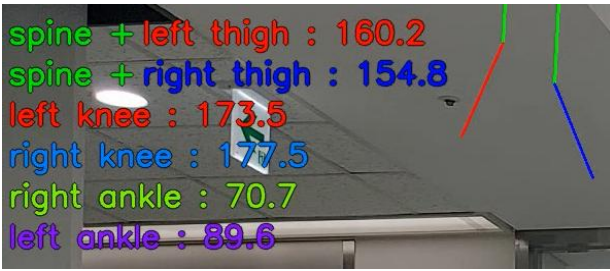


그림 11. 측면 영상 내 분석 표시
Fig. 11. Analysis display of sagittal image

$$\theta_{joint} = \text{acos}\left(\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\| \|\mathbf{v}_2\|}\right) \times \frac{180}{\pi}. \quad (4)$$

다른 관절의 각도를 구하는 경우에는 v_1 의 각도를 구성하는 벡터 중 첫 번째 벡터, v_2 의 각도를 구성하는 벡터 중 두 번째 벡터를 대입하면 된다. 발목의 각도의 경우에 발목 위치에서 각도를 구할 경우 관절의 위치상에서 발생하는 각도에 의해 정확한 발목 각도를 계산할 수 없기에 그림 10과 같이 발꿈치에서 엄지발가락을 이은 직선 벡터를 통하여 구하였다.

계산 후 나온 각도를 분석 영상 왼쪽 상단에 OpenCV를 이용하여 나타냈으며, 그림 11과 같이 허벅지와 척추의 각도는 직관적으로 확인 가능하게 그려 넣었다. 또한 각 관절에서 구해진 각도를 나타내었다.

III. 실험 결과

1. 사용한 보행영상 데이터 셋

뇌성마비 소아환자의 데이터가 부족한 관계로 정상인 영상으로만 제한한 시스템의 성능을 테스트한다. 사용한 테스트 데이터 셋은 정면 영상 5개, 측면 영상 3개로 총 8개의 영상을 사용하였다. 정면 측면 공통으로 고정된 카메라를 이용하여 촬영하였으며, 정면의 경우 그림 12와 같이 완전히 정면에서, 측면의 경우 그림 13과 같이 완전히 측면에서 촬영하였다.

기울기 모델은 MPII human pose dataset [12], World Cities dataset [13]을 데이터 셋으로 사용하였다. MPII



그림 12. 정면 영상 예시
Fig. 12. Frontal image example



그림 13. 측면 영상 예시
Fig. 13. Sagittal image example

human pose dataset은 약 4만 명의 인물이 포함된 2만 5천장의 데이터 셋으로 유튜브 영상에서 추출되었다. 2만 5천장 중에서 기울어있는 영상을 제외하고 5,022장만 데이터 셋으로 사용하였다. World Cities dataset은 Flickr 웹사이트로부터 40개의 주요 도시들의 지리학적 쿼리를 이용하여 수집한 22,994장으로 구성되어 있으며, 사람이 적거나 없는 기울어있지 않은 영상만 수집하여 2,207장만 데이터 셋으로 사용하였다. 두 가지 데이터 셋을 합쳐서 총 7,229장을 최종적인 학습용 데이터 셋으로 사용하였다.

2. 영상 분석 성능 평가

2.1 평가 지표

정면 영상에서 관절의 좌표는 MPJPE (Mean Per Joint Position Error), 어깨와 골반의 불균형 각도는 MAE (Mean Absolute Error)를 통해서 평가하고 단위는 각각 픽셀과 도 (°)이며, 측면 영상에서 관절의 좌표는 MPJPE, 관절 각도 MAE를 통해 평가하고 단위는 각각 픽셀과 도 (°)이다.

$$MAE = \frac{1}{V} \sum_{v=1}^V MAE_v, \quad (5)$$

$$MAE_v = \frac{1}{N} \sum_n \sum_{t=1}^T |\hat{\theta}_{nt} - \theta_{nt}|.$$

$$S_v = \sum_{t=1}^T \sum_{i=1}^N \|(J_i^{(t)} - J_r^{(t)}) - (\hat{J}_i^{(t)} - \hat{J}_r^{(t)})\|, \quad (6)$$

$$S = \sum_{v=1}^V S_v, \quad MPJPE = \frac{S}{T \times N \times V}.$$

표 1. 정면 영상에서 시스템 성능 비교

Table 1. Comparison of system performance in frontal videos

	Shoulder	Pelvis	Average	MPJPE
Original	4.172°	4.220°	4.196°	8.652
Pre	4.096°	3.871°	3.984°	8.746
Post	3.048°	2.985°	3.017°	8.699
Pre+Post	3.077°	2.865°	2.971°	8.514

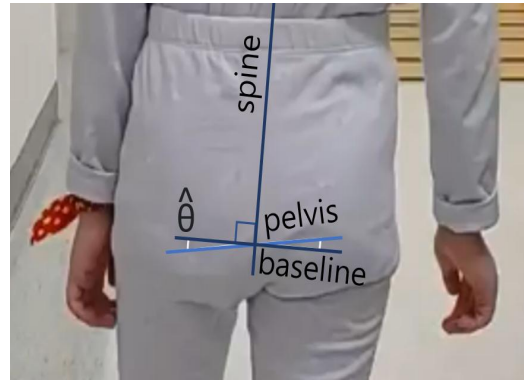
MAE는 수식 (5), MPJPE은 수식 (6)을 통해 계산된다. MAE는 각 관절에서의 실제 값과 측정값과의 차이의 평균으로, 수식 (5)에서 V 는 영상의 개수, v 는 현재의 영상번호, MAE_v 는 v 번째 영상의 MAE, N 은 대칭인 관절의 개수로 본 논문에서는 좌우 관절을 계산하기에 $N=2$ 로 사용하며, n 은 현재 관절의 번호, T 는 전체 프레임의 수, t 는 현재 프레임 번호, $\hat{\theta}_{nt}$ 은 현재 프레임에서의 현재 관절의 예측값, θ_{nt} 는 현재 프레임에서의 현재 관절의 GT값을 나타낸다. MPJPE는 모든 관절의 추정 좌표와 실제 좌표 거리의 평균을 구하여 산출되는 지표로 수식 (6)에서 S_v 은 해당 영상에서의 오차의 합, S 는 전체 영상에서의 오차의 합, V 는 영상의 개수, v 는 현재의 영상번호, T 는 전체 프레임의 수, N 은 관절의 전체 개수, J_r 은 기준이 되는 중심 관절로 본 논문에서는 18번 관절을 J_r 로 지정하였다. 정면 영상은 뇌성마비 소아 환자와 정상인을 구분해서 평가한다. GT는 각각의 영상에 직접 라벨링 하여 제작하였다.

2.2 분석 시스템 평가

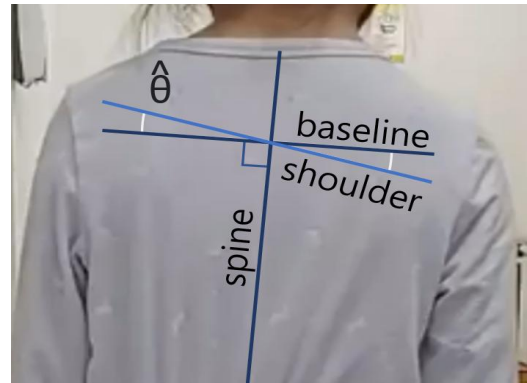
제안한 시스템을 평가하기 위해 4가지 시스템에 대한 성능을 비교한다. 이는 각각 original, pre, post, pre+post로 original은 전/후처리를 하지 않고 자세 추정 모델과 불균형 탐지 알고리즘만을 이용한 시스템, pre는 original에 전처리(영상 기술기 보정)를 적용한 시스템, post는 original에 후처리(temporal smoothing)를 적용한 시스템, 마지막으로 pre+post는 모든 제안한 전처리와 후처리를 적용한 시스템이다. 각 지표의 계산에 있어서 측면, 정면 영상들에서 계산된 MAE와 MPJPE들을 측면, 정면 영상별로 평균을 계산하여 최종적인 지표를 구한다.

정면 영상에서의 평가 요소는 불균형 각도의 MAE와 MPJPE를 사용하며 각 불균형 각도의 MAE는 오른쪽 각도와 왼쪽 각도를 포함한 계산이다. 불균형 각도에서 어깨는 어깨 불균형 각도의 MAE, 골반은 골반의 불균형 각도 MAE, 평균은 어깨와 골반 MAE의 평균을 나타내며, 표 1에서는 각각 shoulder, pelvis, average로 나타난다.

각 관절의 계산에 있어서, 골반 불균형 각도를 계산할 때 사용하는 골반의 불균형 각도는 그림 14-(a)와 같이 기준선과 골반 연결선 사이의 각도이며, 골반 연결선이 기준선 위로 올라가는 경우의 각도는 양수이다. 어깨 불균형 각도를 계산할 때 사용하는 어깨의 불균형 각도 또한 그림 14-(b)



(a)



(b)

그림 14. 정면 영상에서의 분석 영상 : (a) 골반에서의 기준선과 골반 연결선, (b) 척추선, 어깨에서의 기준선과 어깨 연결선

Fig. 14. Analysis image of frontal image : (a) reference line of pelvis and pelvis line, (b) Spine line, reference line of shoulder and shoulder line

와 같이 가로로 놓인 기준선과 어깨 연결선 사이의 각도이다. 이렇게 얻은 불균형 각도들을, 각각의 GT를 통해 구해진 불균형 각도들과 비교하여 MAE를 계산한다.

측면 영상에서의 평가 요소로 관절 각도의 MAE와 MPJPE를 사용하며, MAE의 경우 왼쪽과 오른쪽 관절 각도를 모두 각각 구하여 얻은 MAE의 평균값을 사용한다. 표 2에서 허벅지, 무릎, 발목 및 이들의 평균은 각각 thigh, knee, ankle, average로 표시되며, 이들은 허벅지와 척추 간의 각도, 무릎의 각도, 발목의 각도에 대한 MAE와 이 세 항목의 평균을 나타낸다.

표 2. 측면 영상에서 시스템 성능 비교

Table 2. Comparison of system performance in sagittal videos

	Thigh	Knee	Ankle	Average	MPJPE
Original	4.870°	6.118°	6.678°	5.889°	7.902
Pre	4.828°	6.468°	7.218°	6.171°	7.836
Post	5.238°	8.115°	6.667°	6.673°	6.704
Pre+Post	4.663°	5.961°	5.703°	5.442°	6.899



(a) (b) (c)

그림 15. 측면 영상에서의 분석 영상 : (a) 척추(초록색)와 허벅지(왼쪽 : 빨간색, 오른쪽 : 파란색), (b) 오른쪽 무릎(허벅지 : 파란색, 종아리 : 연두색), (c) 오른쪽 발목

Fig. 15. Analysis image of sagittal image : (a) spine (green) and thigh (left : red, right : blue), (b) right knee (thigh : blue, calf : light green), (c) right ankle

각 관절 각도 계산에서 허벅지 관절 각도는 그림 15-(a)에서와 같이 척추 연결선, 왼쪽 및 오른쪽 허벅지 연결선과의 각도를 사용한다. 무릎 관절 각도는 그림 15-(b)에서와 같이 한쪽 종아리 연결선과 같은 쪽 허벅지 연결선 사이의 각도를 사용한다. 마지막으로, 발목 관절 각도는 그림 15-(c)에서와 같이 허벅지 연결선과 발 연결선 사이의 각도를 사용한다. 이렇게 얻은 각 관절의 각도들을, 각각의 GT를 통해 구해진 관절의 각도들과 비교하여 MAE를 계산한다.

표 1은 정면 영상에서 시스템 성능 비교로, pre+post에서 pelvis, average의 불균형 각도 MAE가 2.865°, 2.971°로 가장 좋은 모습을 보여주고, post에서는 shoulder의 불균형 각도 MAE가 3.048°로 가장 높은 성능을 보인다. 이때 original은 shoulder, pelvis, average의 불균형 각도 MAE가 4.172°, 4.220°, 4.196°인 것에 비해 pre+post의 성능은 각각 1.095°, 1.355°, 1.225°만큼 MAE가 감소한 것을 볼 수 있다. 또한 MPJPE도 original에 비해서 pre+post가 0.138픽셀만큼 더 높은 성능을 보인다.

표 2는 측면 영상에서 시스템 성능 비교로, pre+post에서 thigh, knee, ankle, average 각도의 MAE가 4.663°, 5.961°, 5.703°, 5.442°로 가장 높은 성능을 보여준다. 이때 original은 thigh, knee, ankle, average 각도의 MAE가 4.870°, 6.118°, 6.678°, 5.889°인 것에 비해 pre+post의 각도의 MAE는 각각 0.207°, 0.157°, 0.975°, 0.447°만큼 MAE가 감소한 것을 볼 수 있다.

표 2의 pre와 post를 보면 original보다 성능의 향상이 약간 있거나 변화가 거의 없으나 pre+post에서 최종적으로 성능이 더욱 향상됨을 볼 수 있다. 표 1에서도 pre, post에 비해서 pre+post가 더 나은 성능임을 볼 수 있다. 이를 통해 전처리와 후처리를 같이 사용하는 경우가 가장 좋은 성능을 도출할 수 있음을 알 수 있다.

최종적인 출력 결과는 정면 영상의 경우 그림 9와 같이 어깨와 골반의 불균형 각도를 시각화하여 나타내고, 그림 16과 같이 OpenCV를 이용하여 대상의 몸에 척추, 어깨와 골반을 나타내는 선을 각각 그려 넣고, 어깨와 골반의 기준선



그림 16. 정면 영상의 보행 분석 시스템 결과 예시
Fig. 16. Demonstration of the gait analysis system on frontal video



그림 17. 측면 영상의 보행 분석 시스템 결과 예시
Fig. 17. Demonstration of the gait analysis system on sagittal video

또한 그려 넣었다. 측면 영상의 경우, 왼쪽 상단에 그림 11과 같이 척추와 허벅지, 무릎, 발목의 각도를 시각화하여 나타내고, 그림 17과 같이 대상의 몸에 척추, 허벅지, 종아리, 발을 나타내는 선을 그려 넣었다. 이를 통하여 최종적으로 나온 영상에서 분석한 대상의 걸음걸이를 한눈에 볼 수 있게 하였다.

IV. 결론

1. 실험 결론

본 논문에서는 스마트폰 동영상을 활용한 자세 불균형 탐지를 위한 스마트폰 동영상 기반 보행 분석 시스템을 제안한다. 보행 분석을 수행하기 위해서 human pose estimation 모델인 AlphaPose를 활용하여 보행 중인 사람의 관절의 위

치를 추정하였다. 이때 전처리로서 기울기 보정 모델을 사용하여 스마트폰 영상에서의 human pose estimation을 더 용이하게 하였고, 후처리로서 SmoothNet을 적용하여 human pose estimation에서 발생하는 jitter를 줄여서 분석에 용이하게 하였다. 그리고 전처리와 후처리의 적용 여부에 따른 시스템 분석 성능을 정면 영상과 측면 영상에 대해 비교하였다. 스마트폰 영상을 전처리와 후처리로 보정하여 더 나은 분석을 할 수 있게 제시한 방법은 오차율이 정면 영상에서는 4.196°에서 2.971°, 측면 영상에서는 5.889°에서 5.442°로 가장 좋은 성능을 보여주었다. 이를 통해 다른 특별한 장비 없이 스마트폰 동영상만을 이용해서 의료 실무에서 사용할 수 있는 지표들 (어깨 불균형 각도, 척추와 허벅지의 각도, 무릎 각도, 발목 각도)을 높은 정확도로 분석하고, 자세 불균형 탐지가 가능한 시스템을 구축하였다.

2. 한계점

SmoothNet의 과도한 jitter 제거도 문제이다. 원본 영상에서 추출한 관절 좌표에는 jitter가 많이 섞여 있지만, SmoothNet은 해당 jitter 말고도 특징 자체를 없애는 경우도 발생했다. 이를 보완할 방법에 관한 후속 연구를 진행할 예정이다.

V. References

- [1] A. Mirelman, P. Bonato, R. Camicioli, T. D. Ellis, N. Giladi, J. L. Hamilton, C. J. Hass, J. M. Hausdorff, E. Pelosin, Q.J. Almeida, "Gait impairments in Parkinson's Disease," *The Lancet Neurology*, Vol. 18, No. 7, pp. 697-708, 2019.
- [2] S. Hwang, S. Ha, "A Study on Parents Involvement in Rehabilitation Treatments for Children with Cerebral Palsy," *Journal of Rehabilitation Research*, Vol. 3, No. 17, pp. 309-328, 2013 (in Korean).
- [3] G. J. Lee, T. S. Lee, "Trends in the Use of Artificial Intelligence in Medical Image Analysis," *Journal of the Korean Society of Radiology*, Vol. 16, No. 4, pp. 453-462, 2022 (in Korean).
- [4] D. Qiao, F. Zulkernine, R. Masroor, R. Rasool, N. Jaffar, "Measuring Heart Rate and Heart Rate Variability with Smartphone Camera," In *Proceedings of 2021 22nd IEEE International Conference on Mobile Data Management*, pp. 248-249, 2021.
- [5] E. Hong, S. Cho, S. Lee, J. Jeon, "Deep Learning based Photo Horizon Correction," *Journal of the Korea Computer Graphics Society*, Vol. 23, No. 3, pp. 95-103, 2017.
- [6] H. S. Fang, J. Li, H. Tang, C. Xu, H. Zhu, Y. Xiu, Y. L. Li, C. Lu, "AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
- [7] A. Zeng, L. Yang, X. Ju, J. Li, J. Wang, Q. Xu, "Smoothnet: a Plug-and-play Network for Refining Human Poses in Videos," in *Proceedings of European Conference on Computer Vision*, pp. 625-642, 2022.
- [8] Y. Chen, Y. Tian, M. He, "Monocular Human Pose Estimation: A Survey of Deep Learning-based Methods," *Computer Vision and Image Understanding*, Vol. 192, No. 102897, 2020.
- [9] Z. Cao, T. Simon, S. E. Wei, Y. Sheikh, "Realtime Multi-person 2d Pose Estimation Using Part Affinity Fields," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7291-7299, 2017.
- [10] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick, "Microsoft coco: Common Objects in Context," In *Proceedings of European Conference on Computer Vision*, pp. 740-755, 2014.
- [11] J. H. Shin, R. Yu, J. N. Ong, C. Y. Lee, S. H. Jeon, H. Park, H. J. Kim, J. Lee, B. Jeon, "Quantitative Gait Analysis Using a Pose-estimation Algorithm with a Single 2D-video of Parkinson's Disease Patients," *Journal of Parkinson's Disease*, Vol. 11, No. 3, pp. 1271-1283, 2021.
- [12] M. Andriluka, L. Pishchulin, P. Gehler, B. Schiele, "2d Human Pose Estimation: New Benchmark and State of the Art Analysis," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3686-3693, 2014.
- [13] G. Toliás, Y. Avrithis, "Speeded-up, Relaxed Spatial Matching," In *Proceedings of International Conference on Computer Vision*, pp. 1653-1660, 2011.
- [14] S. Jin, L. Xu, J. Xu, C. Wang, W. Liu, C. Qian, W. Ouyang, P. Luo, "Whole-body Human Pose Estimation in the Wild," In *Proceedings of European Conference on Computer Vision*, pp. 196-214, 2020.
- [15] C. Ionescu, D. Papava, V. Olaru, C. Sminchisescu, "Human3.6m: Large Scale Datasets and Predictive Methods for 3d Human Sensing in Natural Environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 36, No. 7, pp. 1325-1339, 2013.
- [16] K. Sun, B. Xiao, D. Liu, J. Wang, "Deep High-resolution Representation Learning for Human Pose Estimation," In *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, pp. 5693-5703, 2019.
- [17] J. Li, S. Bian, A. Zeng, C. Wang, B. Pang, W. Liu, C. Lu, "Human Pose Regression with Residual Log-likelihood Estimation," In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 11025-11034, 2021.

Yoonho Hwang (황 윤 호)

2018~Computer Science and Engineering
from Kyungpook National University
(B.S.).

Field of Interests: Computer Vision, Human Pose Estimation
Email: ghkddbsh99@knu.ac.kr

Sanghyeon Lee (이 상 현)

2022 Mathematics from Kyungpook
National University (B.S.).
2022~Computer Science and Engineering
from Kyungpook National University
(M.S.).

Field of Interests: Computer Vision, Human Pose Estimation
Email: hyeon1263@knu.ac.kr

Yu-Sun Min (민 유 선)

2006 Seoul National University College of
Medicine (B.S.)
2012 Seoul National University College of
Medicine (M.S.)
2016~Department of Rehabilitation
Medicine, Kyungpook National
University School of Medicine
(Assoc. Prof.)

Field of Interests: Pediatric rehabilitation and exercise science
Email: ssuni119@gmail.com

Jong Taek Lee (이 종 택)

2005 Electrical Engineering from Korea
Advanced Institute of Science and
Technology (B.S.)
2007 Electrical and Computer Engineering
from the University of Texas at
Austin (M.S.)

2012 Electrical and Computer Engineering from the University
of Texas at Austin (Ph.D.)
2012~2022 Electronics and Telecommunications Research
Institute (Senior Researcher)
2022~School of Computer Science and Engineering from
Kyungpook National University (Assist. Prof.)
Field of Interests: Computer Vision, traffic surveillance,
rehabilitation AI
Email: jongtaeklee@knu.ac.kr