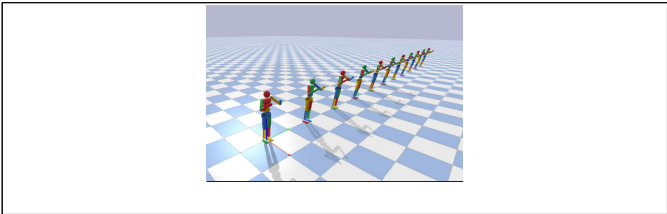


2025학년도 대구과학고등학교
사이언키피아 연구보고서

인간 걷기 데이터를 활용한 PPO와 유전 알고리즘 비교: DTW를 활용한 유사도 및 성능 분석

탐구 영역	정보
-------	----



2025. 06. 02.

학년	반	번호	이름
2	3	3	김동욱
2	5	12	이지현

1. 연구동기 및 목적

가. 연구동기

최근 컴퓨터 비전과 인공지능 기술의 발전으로 OpenPose와 같은 영상 기반 인간 행동 인식 도구가 보행 분석 연구에서 활발히 활용되고 있다. 특히 인간의 보행 동작은 휴머노이드 로봇 및 가상 캐릭터의 자연스러운 움직임을 구현하는 데 핵심적인 요소이다. 그러나 기존의 방법들은 주로 모션 캡처 장비를 사용하여 데이터를 수집하거나, 인공지능 모델을 통해 학습된 제한된 움직임 데이터를 활용하는 데 머물러 있었다. 이에 본 연구에서는 접근성이 뛰어난 OpenPose를 활용하여 실제 사람의 걷기 동작을 쉽고 효율적으로 수집하고, 다양한 보행 패턴 분석 기법을 탐색하여 보다 실질적이고 정확한 동작 모델링 방법을 제안하고자 한다.



나. 연구목적

1) 강화 학습은 최근 로봇틱스 및 가상 환경에서 인간과 유사한 자연스러운 움직임을 구현하는데 널리 사용되고 있으며, 특히 Proximal Policy Optimization(PPO) 알고리즘이 대표적인 예이다. 반면 진화 알고리즘 중 유전 알고리즘(GA)은 다양한 움직임 변이를 적용하여 최적의 동작 패턴을 탐색하는 데 효과적이라고 알려져 있다. 하지만 PPO와 GA가 실제 사람의 걷기 데이터를 기반으로 얼마나 자연스럽게 효율적인 움직임을 생성하는지에 대한 비교 연구는 여전히 부족한 실정이다. 이에 본 연구의 목적은 OpenPose를 통해 수집한 실제 걷기 데이터를 기반으로 PPO와 GA를 적용하고, 각 알고리즘이 생성하는 보행 패턴의 자연성 및 효율성을 정량적으로 비교 분석하는 것이다. 궁극적으로 가장 인간과 유사한 보행 패턴을 효율적으로 학습하고 생성하는 데 적합한 알고리즘을 밝히고자 한다.

2. 이론적 배경

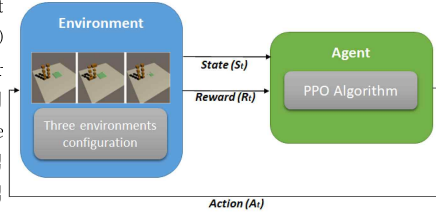
가. OpenPose

1) OpenPose는 심층 신경망(Deep Neural Network)을 기반으로 인간의 신체 키포인트(keypoint)를 실시간으로 추출하는 오픈 소스 컴퓨터 비전 소프트웨어이다. 구체적으로는 Convolutional Neural Network(CNN)을 사용하여 이미지에서 각 신체 부위의 위치를 예측하고 이를 관절 좌표 데이터로 변환한다. OpenPose의 핵심 구조는 Part Affinity Fields(PAFs)를 통해 개별 신체 키포인트들을 효율적으로 연결하며, 높은 정확도와 신뢰성으로 걷기와 같은 동적 움직임 분석에 널리 사용된다. 본 연구에서는 OpenPose를 활용하여 보행 데이터를 수집하고, 추출된 데이터를 분석 가능한 형태로 변환하여 시뮬레이션

환경에 적용한다.

나. Proximal Policy Optimization (PPO)

1) PPO는 심층 강화 학습(Deep Reinforcement Learning, DRL) 알고리즘 중 하나로, 정책(policy)의 업데이트 폭을 제한하여 안정적인 학습을 보장한다. PPO는 Actor-Critic 구조를 가지며, 정책 네트워크(policy network)와 가치 네트워크(value network)를 동시에 학습시켜 정책의 성능을 지속적으로 개선한다. PPO는 기존의 정책 그래디언트 방법에서 클리핑(clipping) 메커니즘을 도입하여 정책 업데이트를 신중히 수행함으로써 학습 안정성을 높인다. 본 연구에서는 PPO를 통해 휴머노이드 로봇이 수집된 인간 보행 데이터를 바탕으로 인간과 유사한 걷기 동작을 학습하도록 설계한다.



다. 유전 알고리즘 (Genetic Algorithm, GA)

1) 유전 알고리즘 (Genetic Algorithm, GA)은 진화 연산 기반의 최적화 방법으로, 자연의 진화 과정에서 영감을 받은 선택(selection), 교배(crossover), 변이(mutation) 과정을 통해 최적해를 탐색한다. 유전 알고리즘은 다음의 단계로 구성된다:

초기 집단(initial population) 생성

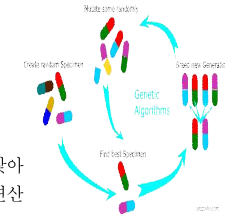
적합도 평가(fitness evaluation)

선택(selection): 적합도에 따라 우수한 개체 선택

교배(crossover): 선택된 개체들 간의 유전 정보 교환

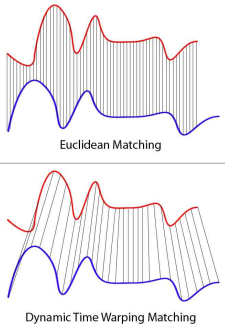
변이(mutation): 일정 확률로 무작위 변이를 가하여 다양성 유지

이 과정이 반복적으로 수행되어 보행 동작 패턴 중 가장 자연스러운 형태를 찾아낸다. 본 연구에서는 휴머노이드의 보행 패턴을 유전자로 정의하고, GA의 연산을 통해 최적의 걷기 동작을 탐색한다.



라. Dynamic Time Warping (DTW)

1) DTW는 시간적으로 비선형적인 특성을 가진 두 시계열 데이터의 유사도를 정량적으로 측정하는 알고리즘이다. DTW의 핵심 원리는 두 시계열 데이터를 비교할 때, 한 시퀀스를 다른 시퀀스에 맞추어 늘리거나 압축하는 방식으로 시간적 불일치를 해결한다. 이를 통해 데이터의 길이나 주기가 다르더라도 정확한 유사도 평가가 가능하다. 본 연구에서는 DTW를 사용하여 PPO 및 유전 알고리즘이 생성한 걷기 패턴과 실제 인간의 걷기 데이터를 비교하고, 이 결과를 통해 알고리즘의 성능과 효율성을 평가한다.



3. 연구 내용

가. 데이터 수집 및 분석 개요 (데이터 수집 개요)

본 연구는 인간의 걷기 동작을 정확히 학습하고 평가하기 위해 다음과 같은 구체적인 방법으로 데이터를 수집하고 분석하였다.

나. 데이터 수집 (정면 및 측면 영상 촬영)

데이터의 신뢰성과 정확성을 높이기 위해 두 대의 스마트폰 카메라를 사용하여 피험자의 걷기 동작을 정면과 측면 두 방향에서 동시에 촬영하였다. DroidCam 소프트웨어를 활용하여 실시간으로 촬영된 영상은 즉시 컴퓨터로 전송되었다. 양방향 촬영은 보행 시 발생하는 전후 및 좌우의 미세한 움직임과 자세 변화를 모두 기록하여 보다 완전한 보행 데이터를 확보하기 위함이다.

다. 데이터 전처리 및 변환 (OpenPose 기반 관절 각도 추출)

촬영된 영상은 OpenPose를 통해 실시간으로 분석되어 신체의 주요 관절 키포인트가 추출되었다. 이후, 각 관절의 좌표를 3차원 형태로 변환한 후, PyBullet 시뮬레이션 환경에 적합한 관절 각도 데이터 형태로 구성하였다. 전처리 단계에서는 데이터의 노이즈 제거 및 좌표 보정 작업을 수행하여 시뮬레이션에서 더욱 정확한 동작을 구현할 수 있도록 하였다.

라. 시뮬레이션 환경 구축 (PyBullet 기반 동역학 모델 구현)

PyBullet을 이용하여 휴머노이드 로봇 모델을 생성하였으며, 전처리된 관절 각도 데이터를 초기 동작값으로 설정하였다. PyBullet 환경에서는 중력, 마찰력, 관절 구속 조건 등 실제 물리적 환경과 유사한 조건을 설정하여 현실성을 높였다.

마. PPO 및 유전 알고리즘 학습 (보상 함수 및 적합도 기반 학습)

PPO와 유전 알고리즘을 개별적으로 적용하여 걷기 동작을 학습하였다. PPO는 보상 함수로 x축 방향 이동 거리를 최대화하며, GA는 동일한 기준으로 적합도를 평가한다. 두 경우 모두 로봇의 머리 z좌표가 1.0 이하가 되면 넘어진 것으로 간주하고 시뮬레이션을 종료하며 페널티를 부여한다.

바. 유사도 평가 (DTW를 통한 알고리즘 성능 분석)

학습 결과로 생성된 보행 동작은 DTW를 통해 실제 인간 보행 데이터와 비교되었으며, 이를 통해 두 알고리즘의 성능과 유사도를 정량적으로 분석하였다. 학습 완료 후, DTW 알고리즘을 이용하여 각 알고리즘이 생성한 보행 패턴과 원본 인간 보행 데이터를 비교하였다. DTW의 유사도 측정 결과를 바탕으로 PPO와 유전 알고리즘의 학습 성능과 효율성을 객관적이고 정량적으로 평가하였다.

4. 연구 방법 및 과정

가. 실험 설계 및 전체 구조 (연구 흐름 개요)

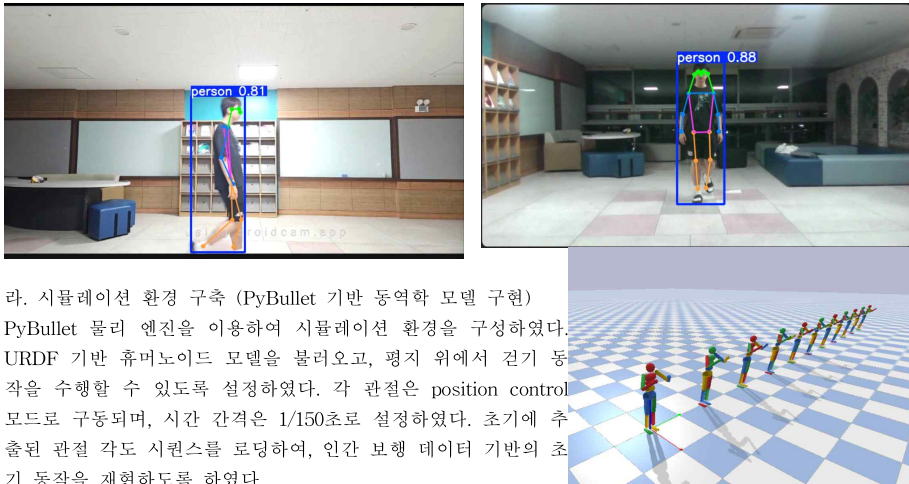
나. 데이터 수집 (정면 및 측면 영상 촬영)

실험자는 피험자 1명을 대상으로 실내 공간에서 3m 정도의 직선 경로를 걷도록 하였다. 두 대의 스마트폰은 각각 피험자의 전면과 측면이 모두 보이도록 90도 각도로 배치되었으며, DroidCam을 통해 두 방향 영상을 동시에 컴퓨터로 전송 및 저장하였다. 촬영은 30fps로 고정되었으며, 전체 약 4초 분량의 동작이 수집되었다. 이는 보행 주기 2회에 해당하며, 각 주기 내 관절의 반복적 변화를 반영한다.



다. 데이터 전처리 및 변환 (OpenPose 기반 관절 각도 추출)

수집된 영상은 OpenPose를 통해 프레임별로 분석되었고, 각 관절에 해당하는 키포인트가 2D 좌표로 추출되었다. 전면 영상은 좌우 관절의 수평 움직임과 팔과 다리가 어깨와 평행한 방향으로 움직이는 동작을 확인하는 데 사용되었고, 측면 영상은 다리와 어깨, 팔꿈치 등의 스윙과 같이 보행의 전후 움직임을 보완하였다. 두 데이터를 결합하여 관절 간 상대 위치 벡터를 계산하고, 이로부터 각 관절의 회전 각도를 계산하였다. 이후 PyBullet에서 제어 가능한 관절 인덱스에 대응되도록 정리한 후, 넘파이 배열 형태로 저장하였다.



라. 시뮬레이션 환경 구축 (PyBullet 기반 동역학 모델 구현)

PyBullet 물리 엔진을 이용하여 시뮬레이션 환경을 구성하였다. URDF 기반 휴머노이드 모델을 불러오고, 평지 위에서 걷기 동작을 수행할 수 있도록 설정하였다. 각 관절은 position control 모드로 구동되며, 시간 간격은 1/150초로 설정하였다. 초기에 추출된 관절 각도 시퀀스를 로딩하여, 인간 보행 데이터 기반의 초기 동작을 재현하도록 하였다.

마. PPO 및 유전 알고리즘 학습 (보상 함수 및 적합도 기반 학습)

PPO는 학습 에이전트가 일정 시간 동안 보행을 수행하며 보상으로 x축 방향의 전진 거리를 최대화하도록 훈련된다. Actor-Critic 구조에서 정책 네트워크는 연속적인 관절 제어 값을 출력하며, 메 에피소드마다 초기 관절 각도 및 환경 잠음이 변화한다. 보상 함수는 전진 거리 기준으로 하되, 머리의 z좌표가 1.0 이하로 떨어질 경우 즉시 시뮬레이션을 종료하므로써 페널티를 가한다.

유전 알고리즘에서는 각 개체가 특정한 관절 각도 배열(유전자)을 가지며, 시뮬레이션 내에서 해당 각도를 반복 적용함으로써 보행을 수행한다. 적합도는 이동 거리로 측정되며, 넘어짐 조건은 PPO와 동일하다. 선택은 가장 적합도가 높은 두 개체를 부모 개체로, 다음 세대에 포함시키고, 교배는 균등 교차 방식으로, 변이는 15% 확률로 조정한다. 총 1000세대를 반복하며 가장 우수한 보행 시퀀스를 도출하였다.

바. 유사도 평가 (DTW를 통한 알고리즘 성능 분석)

알고리즘별로 생성된 관절 각도 시퀀스를 원본 데이터와 비교하기 위해 DTW를 적용하였다. 좌우 다리의 고관절과 무릎, 팔의 어깨 관절 등 총 8개의 주요 관절에 대해 시간축에 따른 유사도를 개별적으로 측정한 후 평균을 취하였다. DTW 결과는 거리값으로 산출되며, 값이 작을수록 원본과 유사함을 의미한다. 이를 통해 PPO와 GA의 보행 동작 유사성을 정량적으로 비교하였다.

학습 결과로 생성된 보행 동작은 DTW를 통해 실제 인간 보행 데이터와 비교되었으며, 이를 통해 두 알고리즘의 성능과 유사도를 정량적으로 분석하였다.

학습 완료 후, DTW 알고리즘을 이용하여 각 알고리즘이 생성한 보행 패턴과 원본 인간 보행 데이터를 비교하였다. DTW의 유사도 측정 결과를 바탕으로 PPO와 유전 알고리즘의 학습 성능과 효율성을 객관적이고 정량적으로 평가하였다.

5. 연구 결과

다음은 작성한 코드이다.

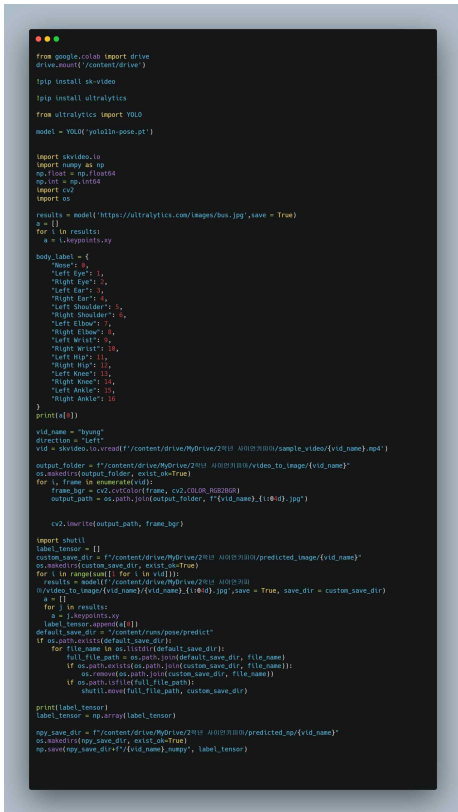


그림 10

영상을 입력받아 프레임 별로 포즈 추측하여 저장하는 코드



그림 12

유전 알고리즘 코드



그림 11

PPO 코드

나. 데이터 수집 (정면 및 측면 영상 촬영)

실험자는 피험자 1명을 대상으로 실내 공간에서 5m 정도의 직선 경로를 왕복 걷도록 하였다. 두 대의 스마트폰은 각각 피험자의 전면과 측면이 모두 보이도록 90도 각도로 배치되었으며, DroidCam을 통해 두 방향 영상을 동시에 컴퓨터로 전송 및 저장하였다. 촬영은 30fps로 고정되었으며, 전체 약 10초 분량의 동작이 수집되었다. 이는 보행 주기 3~5회에 해당하며, 각 주기 내 관절의 반복적 변화를 반영한다.

다. 데이터 전처리 및 변환 (OpenPose 기반 관절 각도 추출)

수집된 영상은 OpenPose를 통해 프레임별로 분석되었고, 각 관절에 해당하는 키포인트가 2D 좌표로 추

출되었다. 전면 영상은 좌우 관절의 수평 움직임과 팔, 다리 스윙 동작을 확인하는 데 사용되었고, 측면 영상은 다리의 굽힘, 골반 회전, 몸통 기울기 등 보행의 전후 움직임을 보완하였다. 두 데이터를 결합하여 관절 간 상대 위치 벡터를 계산하고, 이로부터 각 관절의 회전 각도를 유추하였다. 이후 PyBullet에서 제어 가능한 관절 인덱스에 대응되도록 정리한 후, 넘파이 배열 형태로 저장하였다.

라. 시뮬레이션 환경 구축 (PyBullet 기반 동역학 모델 구현)

PyBullet 물리 엔진을 이용하여 시뮬레이션 환경을 구성하였다. URDF 기반 휴머노이드 모델을 불러오고, 평지 위에서 걷기 동작을 수행할 수 있도록 설정하였다. 각 관절은 position control 모드로 구동되며, 시간 간격은 1/150초로 설정하였다. 초기에 추출된 관절 각도 시퀀스를 로딩하여, 인간 보행 데이터 기반의 초기 동작을 재현하도록 하였다. 바닥면에는 마찰 계수 0.8, 공기 저항은 생략되었으며, 관절 감쇠는 0.05로 조정되었다.

마. PPO 및 유전 알고리즘 학습 (보상 함수 및 적합도 기반 학습)

PPO는 학습 에이전트가 일정 시간 동안 보행을 수행하며 보상으로 x축 방향의 전진 거리를 최대화하도록 훈련된다. Actor-Critic 구조에서 정책 네트워크는 연속적인 관절 제어 값을 출력하며, 매 에피소드마다 초기 관절 각도 및 환경 잡음이 변화한다. 보상 함수는 전진 거리 기준으로 하되, 머리의 z좌표가 1.0 이하로 떨어질 경우 즉시 시뮬레이션을 종료하고 보상을 0으로 만든다.

유전 알고리즘에서는 각 개체가 특정한 관절 각도 배열(유전자)을 가지며, 시뮬레이션 내에서 해당 각도를 반복 적용함으로써 보행을 수행한다. 적합도는 이동 거리로 측정되며, 넘어짐 조건은 PPO와 동일하다. 선택은 상위 30%를 엘리트로 유지하고, 교배는 일점 교차 방식으로, 변이는 10% 확률로 각도 값을 $\pm 10\%$ 범위에서 조정한다. 총 100세대를 반복하며 가장 우수한 보행 시퀀스를 도출하였다.

바. 유사도 평가 (DTW를 통한 알고리즘 성능 분석)

알고리즘별로 생성된 관절 각도 시퀀스를 원본 데이터와 비교하기 위해 DTW를 적용하였다. 좌우 다리의 고관절과 무릎, 팔의 어깨 관절 등 총 8개의 주요 관절에 대해 시간축에 따른 유사도를 개별적으로 측정한 후 평균을 취하였다. DTW 결과는 거리값으로 산출되며, 값이 작을수록 원본과 유사함을 의미한다. 이를 통해 PPO와 GA의 보행 동작 유사성을 정량적으로 비교하였다.

학습 결과로 생성된 보행 동작은 DTW를 통해 실제 인간 보행 데이터와 비교되었으며, 이를 통해 두 알고리즘의 성능과 유사도를 정량적으로 분석하였다.

학습 완료 후, DTW 알고리즘을 이용하여 각 알고리즘이 생성한 보행 패턴과 원본 인간 보행 데이터를 비교하였다. DTW의 유사도 측정 결과를 바탕으로 PPO와 유전 알고리즘의 학습 성능과 효율성을 객관적이고 정량적으로 평가하였다.

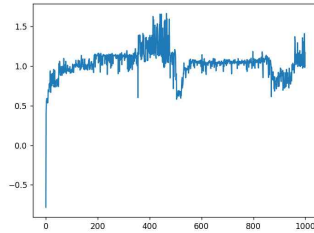
연구 결과 및 분석

본 연구에서는 PPO와 유전 알고리즘(GA)을 동일한 초기 조건과 데이터 기반으로 각각 1000세대(혹은 에피소드)까지 학습시킨 후, 이들의 보행 동작 학습 성능을 정량적으로 비교하였다. 실험 결과는 다음과 같다:

가. 유전 알고리즘의 학습 경향

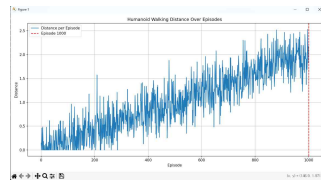
초기 약 50세대까지는 적합도 함수(이동 거리 기준)가 빠르게 증가하여, 초기 해집단에서 효율적인 걷기 동작을 빠르게 찾아가는 경향을 보였다. 그러나 이후에는 적합도가 약 1.0 근처에서 수렴하였으며, 세대가 거듭될수록 큰 성능 향상이 나타나지 않았다. 실제 실험에서는 유전 알고리즘으로 생성된 개체들이 대부분 비정상적인 자세로 시뮬레이션을 시작한 후, 한두 걸음도 걷지 못한 채 전방으로 넘어지는 형태가 반복적으로 관찰되었다. 관찰 결과, 일부 개체는 다리를 뻗는 동작 없이 상체를 빠르게 기울여 앞으로 쓰러지는 방식으로 적합도를 높이기도 하였다. 이러한 현상은 단발적인 움직임으로도 짧은 거리 이동이 가능하여, 단순한 거리 기반 적합도 함수가 구조적으로 오작동할 수 있음을 보여준다. 즉, '실제로

걷는' 움직임이 아니라 '넘어지는 방식의 추진'이 오히려 선택되면서 보행의 자연스러움이나 안정성과는 무관하게 높은 적합도를 얻는 사례가 나타났다. 그 결과, 유전 알고리즘은 걷기 학습에는 실패하였으며, 수렴한 개체들도 모두 자연스러운 보행과는 거리가 먼 결과를 보였다.



나. PPO의 학습 경향

반면 PPO 알고리즘은 초기에 급격한 성능 향상은 없었으나, 학습이 지속될수록 점진적인 성능 향상을 보여주었다. 전체 학습 종료 시점에서는 적합도가 약 2.0 수준까지 도달하였으며, 이는 유전 알고리즘 대비 약 2배 이상의 이동 거리 향상을 의미한다. PPO로 학습된 보행 동작은 완전히 정상적인 걸음이라기보다는 약간 불안정하고 조심스러운 동작이었으며, 종종 자세가 기울어지거나 중심이 흔들리는 양상이 나타났다. 그러나 상체를 수직에 가깝게 유지하며 다리를 번갈아 내딛는 형태로 전진이 이루어졌고, 다리 움직임의 순서와 관절 궤적이 인간의 보행 주기와 유사한 패턴을 보였다. DTW 분석에서도 원본 인간 보행 데이터와의 유사도가 상대적으로 높게 나타났으며, 실제 측정된 평균 DTW 거리는 PPO가 52.7로, 유전 알고리즘의 154.9에 비해 약 3배 가까이 낮은 값을 기록하였다. 이는 PPO가 제한된 관절 제어 조건에서도 보행의 핵심 메커니즘을 어느 정도 학습했음을 정량적으로 입증한다.



다. 비교 분석

학습 속도 측면에서는 유전 알고리즘이 초기 성능 향상에 우위를 보였으나, 최종 보행 성능 및 전진 거리 측면에서는 PPO가 더 우수한 결과를 나타냈다. 이는 정적 최적화 기반의 유전 알고리즘과 동적 정책 기반 학습을 수행하는 PPO의 구조적 차이에 기인한 것으로 해석된다. 또한 PPO는 클리핑 기반 안정화 메커니즘을 통해 불안정한 업데이트를 줄이고, 장기적인 보행 전략을 성공적으로 수립하였다.

6. 연구 결론

본 연구의 목적은 인간 보행 데이터를 기반으로 강화학습(PPO)과 진화 알고리즘(GA)의 보행 동작 학습 성능을 비교하고, DTW를 통해 이들의 유사도 및 효율성을 정량적으로 분석하는 데 있었다. 이를 위해 OpenPose를 활용하여 정면과 측면에서 동시에 보행 영상을 수집하고, 관절 각도를 추출하여 PyBullet 시물

레이션 환경에서 학습 알고리즘에 적용하였다.

유전 알고리즘은 초기 50세대까지 빠르게 적합도를 향상시켰으나, 대부분 비정상적인 보행(예: 상체만 기울여 넘어지는 동작)으로 인해 실질적인 보행을 학습하는 데에는 실패하였다. DTW 유사도는 154.9로, 인간 보행과의 정렬에서 낮은 정확도를 보였다.

반면 PPO는 초기 학습 속도는 느렸으나, 안정적으로 보행 패턴을 개선하며 최종 적합도 약 2.0까지 도달하였다. 생성된 보행은 다리를 번갈아 내딛는 기본적인 주기를 형성하였고, DTW 유사도도 52.7로 상대적으로 높게 나타났다.

결론적으로 본 연구는 강화학습(PPO)이 인간 보행을 모사하는 데 있어 GA보다 구조적으로 더 적합하며, 제한된 관절 제어 환경에서도 시간적 패턴을 효율적으로 학습할 수 있음을 실험적으로 입증하였다. 이러한 결과는 향후 보행 로봇의 제어 전략 설계, 생체모사 인공지능, 가상 인간 동작 생성 등에 있어 실질적인 참고 자료로 활용될 수 있다.

7. 한계점 및 향후 과제

가) 한계점

1. OpenPose 기반 데이터의 한계

본 연구에서 사용된 OpenPose는 2D 이미지 기반의 키포인트 추출 방식이므로, 일부 관절—특히 척추, 허리, 복부 등의 중심부 움직임—에 대한 데이터가 부족하다. 이는 걷기 동작에서 몸통의 균형 유지나 무게 중심의 변화가 중요한 역할을 함에도 불구하고, 학습 알고리즘에 해당 정보를 제공하지 못하게 하여 보행의 안정성 저하로 이어질 수 있다. 또한 OpenPose의 관절 인식 정확도는 피사체의 자세, 조명, 배경 등에 영향을 받으므로, 실험 중 일부 프레임에서는 잘못된 관절 위치가 추출되어 학습 데이터의 품질을 저하시킬 수 있다.

2. 시뮬레이션 환경과 실제 보행의 차이

PyBullet 환경은 현실의 동역학을 근사적으로 모사하지만, 실제 인간의 생체역학과 근골격계의 복잡성을 모두 반영하지 못한다. 특히 관절별 제한, 근육 반작용, 지면 반발력 등의 요소는 단순화되어 있어, 알고리즘이 현실 보행과 완전히 동일한 전략을 학습하기에는 한계가 있다. 이는 DTW 유사도가

높더라도 실제로는 다소 인위적인 동작이 형성될 수 있음을 시사한다.

3. 보상 함수 및 적합도 함수 설계의 구조적 편향

본 연구에서는 보상 및 적합도 기준으로 단일 지표(이동 거리)를 사용하였고, 자세 안정성은 머리 높이 조건으로만 간접 평가되었다. 이로 인해 유전 알고리즘에서는 넘어짐 동작이 오히려 '효율적인 이동 전략'으로 선택되는 구조적 편향이 발생하였다. 이는 향후 보행의 품질을 보다 정교하게 평가하기 위해, 관절 간 에너지 소모, 신체 균형, 주기성 등의 요소를 포함한 다중 지표 기반 평가 기준이 필요함을 시사한다.

4. 강화학습의 일반화 가능성 및 한계

PPO는 실험에서 상대적으로 우수한 보행 패턴을 생성하였으나, 이는 학습 데이터 및 환경 조건에 밀접하게 의존한다. 다른 피험자의 보행 데이터나 외부 환경 조건(예: 경사면, 장애물 등)이 주어질 경우 성능이 유지될 수 있을지는 미지수이다. 강화학습의 일반화 능력을 향상시키기 위해서는 다양한 보행 사례와 복수 환경에서의 반복 학습이 필요하다.

나. 향후 과제

1. 다양한 보행 데이터 수집 및 일반화

본 연구는 단일 피험자의 데이터를 중심으로 진행되었기 때문에, 학습된 보행 모델은 특정 개인의 보행 특성에 과도하게 의존할 수 있다. 향후에는 연령, 성별, 신체 조건이 다양한 다수의 피험자로부터 보행 데이터를 수집하고, 이 데이터를 기반으로 일반화 가능한 보행 제어 알고리즘을 개발할 필요가 있다.

2. 3D 키포인트 및 센서 융합

OpenPose 기반의 2D 키포인트는 신체 일부의 움직임을 반영하지 못하므로, 향후에는 3D 포즈 추정 기법이나 관성 센서(IMU), 모션 캡처(MoCap) 등의 장비를 활용한 다중 센서 융합 기반 데이터 확보가 필요하다. 이를 통해 보다 정밀한 보행 데이터가 확보되고, 신체 중심 이동 및 균형 유지에 대한 정교한 분석이 가능해진다.

3. 보상 함수 고도화 및 다중 목표 최적화

현재 사용된 보상 및 적합도 기준은 단일 전진 거리 기반이므로, 향후에는 에너지 효율성, 보행 안정성, 자세 대칭성 등을 고려한 다중 목표 보상 함수 설계가 필요하다. 이를 통해 실제 인간의 보행 특성과 더욱 유사한 동작이 학습될 수 있다.

4. 복잡한 환경 내 학습 확장

본 연구는 평지 환경을 기반으로 하였으나, 실질적인 응용을 위해서는 경사로, 장애물, 다양한 지면 마찰 조건 등 복잡한 환경에서의 보행 학습 확장이 필요하다. 강화학습 알고리즘이 이러한 외부 변화에 적응할 수 있는지 검증하는 것은 향후 연구의 중요한 방향이 될 수 있다.

<참고 문헌>

1. Villela, L. F. C., & Colombini, E. L. (2017). Humanoid Robot Walking Optimization using Genetic Algorithms. Universidade Estadual de Campinas.

2. Lee, H. S. (2019). Application of Dynamic Time Warping Algorithm for Pattern Similarity of Gait. Journal of Exercise Rehabilitation, 15(4), 526-530.

- 3.