FULL PAPER

A Complex Thinking and Reinforcement Learning Approach for Musculoskeletal Modeling for Prosthetic Control

Adan Dominguez-Ruiz[a], Edgar-Omar López-Caudana[a], Oscar Loyola[b] and Pedro Ponce-Cruz[c]

[a]Institute for the Future of Education, Tecnológico de Monterrey, Mexico; [b] Universidad Autonoma de Chile; [c] Institute for Advanced Manufacturing, Tecnológico de Monterrey, Mexico

ARTICLE HISTORY
Compiled September 18, 2024

ABSTRACT
Musculoskeletal modeling is a vital area of study in biomechanics and rehabilitation engineering, aiming to replicate natural human movements in prosthetic limbs. Current prosthetic control systems often lack adaptability and efficiency, especially in environments that deviate from controlled, flat surfaces. Key variables to assess include joint positions, velocities, muscle activations, and sensory inputs from prosthetic limbs. The absence of effective adaptive systems results in limited functionality, higher energy consumption, and less natural movement patterns in users, particularly in dynamic environments. Addressing these gaps could significantly improve user mobility, comfort, and independence, reducing long-term health complications. Our solution leverages AI-based approaches, combining reinforcement learning with detailed models, which offer better adaptability and decision-making capabilities than traditional control systems. This study expects to achieve a more natural, human-like gait in prosthetic users, enhancing their quality of life. We utilize a novel integration of RL with a highly detailed model for transtibial prostheses, targeting an improvement in gait symmetry, efficiency, and stability. Results indicate the successful implementation of basic walking locomotion, achieving forward movement without falls, with a velocity accuracy of 92.67%. Further research will focus on validating these models with real-world data, setting the stage for more practical applications.

Word Count - 4699 words

KEYWORDS
Musculoskeletal Modeling; Prosthetic Control; Reinforcement Learning; Assistive Robotics; Higher Education

1. Introduction

Development of advanced prosthetic limbs capable of replicating natural human movements is a critical goal in biomechanics and rehabilitation engineering. Traditional prosthetic control systems fall short in terms of adaptability and functionality, limiting the use of the prosthetics to specific places, and separating the movements performed by the prosthesis from the human user [1]. To progress further with recent advances in artificial intelligence (AI), a safe place is required to produce more complex control systems.

CONTACT D. R. Adan. Email: adanydr@tec.mx

The integration of the metaverse, simulation environments, and virtual approaches in research and development is critical at presenting innovation and addressing complex real-world challenges. The use of metaverse technologies allows for creating safe environments, offering unique opportunities for educational innovation and practical applications in various fields[2]. Working in simulation environments before real-world implementation provides a controlled and flexible testing ground that can help optimize strategies, predict outcomes, and reduce potential risks associated with developing new technologies [2,3].

Musculoskeletal (MS) modeling provides a comprehensive framework to understand human motion research, locomotion analysis, and the complex interactions between joints and the movement produced by body muscles and the attachments into the bones. These models simulate the movement and loading of musculoskeletal structures, allowing researchers to analyse how internal and external forces affect the body. The utility of musculoskeletal models in simulating lower extermity's movement can help in the prediction of surgical outcomes and development of rehabilitation strategies [4]. Such models are crucial for exploring biomechanical propierties and movements which otherwise are difficult or impossible to measure directly, providing relevant data for both clinical and research applications [5].

In the context of prosthetic design, a MS model can enable simulation of how prosthetic limbs integrate with the human body, an integration essential for developing prosthetics and control systems which can mimic organic limb movements, enhancing user's mobility and comfort. Some models have shown capabilities to improve gait efficiency and reduce risk of secondary complications like joint degeneration [6], and provided insights into how to optimise prosthetic desgins to better interact with the users remaining musculature and skeletal structure [5,7].

Musculoskeletal modeling helps facilitating the continuous testing and refinement pf prosthetic desgins in a virtual environment, reducing the time associated with physical prototyping, and having certainty about similar results between virtual and physical results [8]. The use of MS models in conjuction with reinforcement learning (RL) algorithms has shown potential at solving highly complex issues involved with humanoid locomotion and predicting human gait, where the agent can learn and improve over time [6,9–11].

RL, as a subfield of AI, has shown great promises in control systems optimization by enabling agents to constantly learn from the observations and interaction with their environment [12]. In prosthetic control, RL algorithms can be used to develop adaptive strategies that respond to the user's movement and environmental changes, constantly improving and adapting, extending limitations of static control systems [13,14]. Current research builds on these advancements by integrating RL with a high-fidelity musculoskeletal model to control a transtibial prosthesis.

Despite the progress made, achieving a human-like walking gait with prosthetic devices remains a significant challenge. Our study addresses this challenge by using a RL-based control system for a virtual MS model with a transtibial prosthesis, consisting of 70 muscles and one electrical actuator at the ankle, resulting in a 71 dimensional action space and 37 degrees of freedom (DoF). The objective is to enable to prosthesis to perform a walking locomotion that is stable and efficient in terms of actuator's activation and effort. The current model successfully achieves forward movement without falling, it does not yet produce a gait that closely ressembles natural human walking. This paper presents our methodology, results, and an analysis of the current limitations and future directions to achieve more realistic prosthetic control, while establishing a proper model for locomotion analysis and more complex control systems development.

The contributions of this study are as follows:

(1) Implementation of a Detailed Musculoskeletal Model with a Transtibial Prosthesis: Development of a virtual MS model consisting of 69 muscles and 1 electrical actuator, with 37 DoF, the study provides a detailed framework that accurately simulates interactions between the model and environment. This level of detail in the illustrates the capability to handle high-dimensional control tasks, commonly found in prosthetic and robotic applications.

(2) Integration of Reinforcement Learning with MS model: This study show the strenghts of using RL with biomechanical simulations to develop advanced control systems for prosthetic limbs, and to predict movement. It shows how AI can be applied to improve adaptability and efficiency of prosthetic devices, crucial in the field of assistive robotics.

(3) Achieving Basic Walking Locomotion: The RL-trained model is able to achieve basic

walking motion without falling, demonstrating the potential of using RL for prosthetics control, showing the AI capability to improve stability and functionality of these devices. Even though human-like gait is not achieved, this paves the way for future improvements.

(4) Foundation for Future Experimental Validation: While current study lacks experimental data, it establishes a robust framework that can be used for future validation with real-world data. This sets the stage for subsequent research that can test the model and refine control systems using experimental results, enhancing its practical applicability and effectiveness.

2. Related Work

In the last 15 years, publications regarding biomechanical models or simulations have increased from 500 per year to more than 2000 yearly, showing the necessity for these musculoskeletal simulated models to study neuromuscular coordination, analyse athletic performance, estimate internal loading of muscles, and even identify sources of pathological movements [15]. Early works on MS modeling were to create dynamic simulations of human movement, while developing OpenSim [4,5,16], creating dynamic simulations of human movement, and demonstrating the utility of MS models in understanding the interactions between muscles, bones, joints, and predicting outcome of some surgical interventions. Models typically include the hip, knee, and ankle koints, fundamental in locomotion [17].

- Hip Join: Producing three degrees of freedom: flexion/extension, abduction/adduction, internal/external rotation ([18]).
- Knee Joint: With a primary single degree of freedom producing flexion/extension, and a small degree of rotation.
- Ankle Joint: Responsible for plantarflexion, dorsiflexion, inversion/eversion to facilitate the ability to adapt to the walking surface ([19,20]).

Representation have been done to study the low limb biomechanics, starting from lower-extremity models with organic limbs, one of the main base models used is provided by OpenSim and designed analysing 21 cadavers and verified by magnetic resonance of 24 young subjects, showing a total of 80 muscles [21]. A detailed analysis of different models in the literature, with number of muscles and DoF, and its primary use is detailed in Table 1, where even basic models with prosthetics were developed for biomechanical studies.

Various simulation platforms have been developed to facilitate musculoskeletal modeling and prosthetic design. Engines as OpenSim, focuses on detailed models of the musculoskeletal system, however, they lack control systems development, and provide limited support for environment interaction, while being computational expensive ([5,27]). A physics simulator such as MuJoCo ([28]) provide training capabilities and efficient environment interactions, but lacks human muscles modelling. Scone, focuses on optimizing control strategies in biomechanics, providing tools for simulating locomotion and understanding the contributions of assistive devices to human gait, by using a very specific programming language (LUA), with limited algorithm development capabilities [29]. A simulation environment such as MyoSuite ([30]) has been developed to cover both deficiencies, allowing the design and testing of control methods to analyse locomotion with musculoskeletal models. This platform is particularly beneficial for prosthetic control, offering the ability to simulate and optimize complex motor tasks and control strategies in a virtual environment before applying them in real-world scenarios.

The integration of Artificial Inteligence (AI), with MS models represent a new approach to solve locomotion issues in the field of biomechanics and prosthetic control. AI-driven techniques allow fort the development of adaptive and intelligent control systems that can simulate and optimise human locomotion in virtual environments, by detecting the surroundings and even decoding intentions [31–33]. Detailed MS models can accurately represent the anatomy and biomechanics of the human body, and with the use of machine learning algorithms, it is possible to manage complex motor tasks, such a walking and balancing, even with the use of passive prostheses, by the use of imitation learning [34]. However, these techniques require a lot of data to generate a neural network model to handle highly actuated systems. RL emerge as a method to predict movements with no use of previous experimental data, capable of imitate natural human walking with organic limbs with the use of convolutional neural networks [10,35,36].

Table 1. Analysis of different musculoskeletal models and its primary purpose

Paper	Number of Muscles	Degrees of Freedom (DoF)	**Purpose**
[4]	44	7	To study orthopaedic surgical procedures and simulate lower extremity movements
[6]	54	23	To optimize human walking dynamics and study gait mechanics
[22]	-	43	DRL control implementation for shared autonomy locomotion
[23]	15	14	To analyse lower-extremity of a human subject with transfemoral amputation wearing a generic lower-limb bone-anchored prosthesis.
[24]	80	20	To run high-fidelity MS simulations integrated with RL for prosthetic control
[5,6]	92 / 54	23	To analyze gait mechanics and muscle function during walking and running
[25]	92	21	Analyse role of muscles in passive and powered prostheses
[26]	38	14	Muscles contribution in walking mechanics (amputees)

Despite the advancements in MS modeling and AI, current prosthetic control systems face several challenges, such as lack of adaptability in real-time to the user's needs, following only stored trajectories [37,38], and suffering from delayed response times at achieving stable movements. The integration of AI techniques in prosthetic control is still in its early stages, and further research is needed to refine these methods and make them more robust and reliable. Addressing these challenges requires the development of advanced control strategies that can adapt to the dynamic requirements of prosthetic users. Our proposed approach emphasizes the integration of MyoSuite with RL algorithms, presenting a solution to the challenges identified in current prosthetic control systems. This approach has the potential to significantly impact the field of prosthetics and wearable technologies, offering enhanced functionality and user experience, broadening the scope to the field of biomechanics and rehabilitation engineering.

Having established the background and identified the gaps in current prosthetic control systems, we now present the methodology used in our study.

3. Methodology

To provide a comprehensive overview of our methodology and the experiments conducted, a flowchart that outlines the sequential steps taken in this study in show in Figure 1, illustrating the progression from the initial development of the musculoskeletal model to the integration with the MyoSuite interface, the implementation of the reinforcement learning algorithm, definition of reward functions, and the experimental setup, leading to the analysis and documentation of the results.

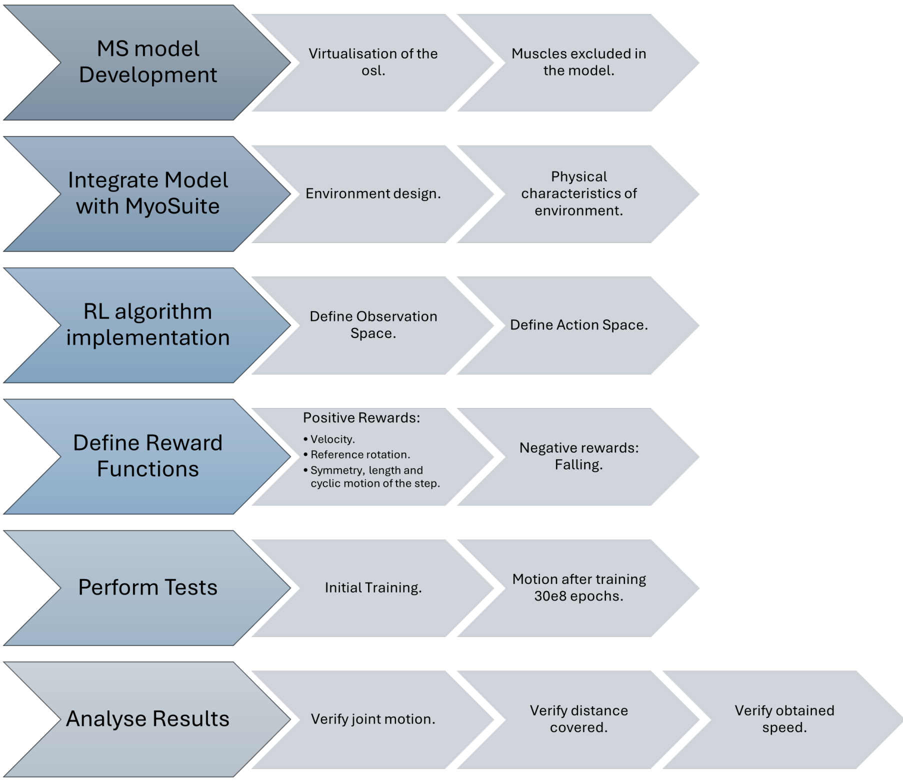

Figure 1. Flowchart of the methodology conducted in this study. The diagram outlines the key steps from the development of the musculoskeletal model, leading to the outcomes of the RL algorithm applied in locomotion.

3.1. Musculoskeletal Model

The MS model developed for this study is a detailed representation of the human lower-extremity body, designed to simulate and study the biomechanics of walking for individuals with transtibial amputations. The model includes 69 muscles, one electrical actuator, and 37 DoF, providing a framework to analyse gait mechanics and optimise prosthetic control strategies. The model is structured to include the anatomical components necessary for realistic movement simulation, such as bones, muscles, and joints, which interact to produce human locomotion patterns.

Muscles in this model are represented using the Hill-type muscle model [5], known for its accuracy in simulating the force-generation properties of real muscle. The Hill-type model consist of three elements: a contractile element, a parallel elastic component and a series of elastic components. In Myosuite, tendons are stiff, so activation patterns are raised across a muscle separated into different parts, with insertion points located in specific parts of the bones, creating a joint movement depending on the activation [30].

Certain muscles that the state-of-the art indicates they do not contribute to lower limb function post-amputation were removed from the model [39,40], such as tibialist anterior, gastrocnemius, and soleus, which are either absent or significantly altered after a transtibial amputation, were excluded. Retained muscles were carefully chosen to ensure the model remains both comprehensive and manageable, allowing for accurate simulations focused on prosthetic control. Details of the muscles involved, their primary functions, joint movements, and whether they were removed on the right leg are summarised in Table 5.

The OpenSourceLegv2 (OSL) prosthesis is a key component of our model [41], offering am modular design that can be adapted for both transfemoral and transtibial amputees. For the presented model, designed for transtibial amputees, the prosthesis extends to the ankle level, ensuring stability and functionality for the lower leg. This modular approach allows for versatility in the design and testing of prosthetic devices, making it possible to tailor the prosthesis to the specific needs of different types of amputations. To adapt the OSL to

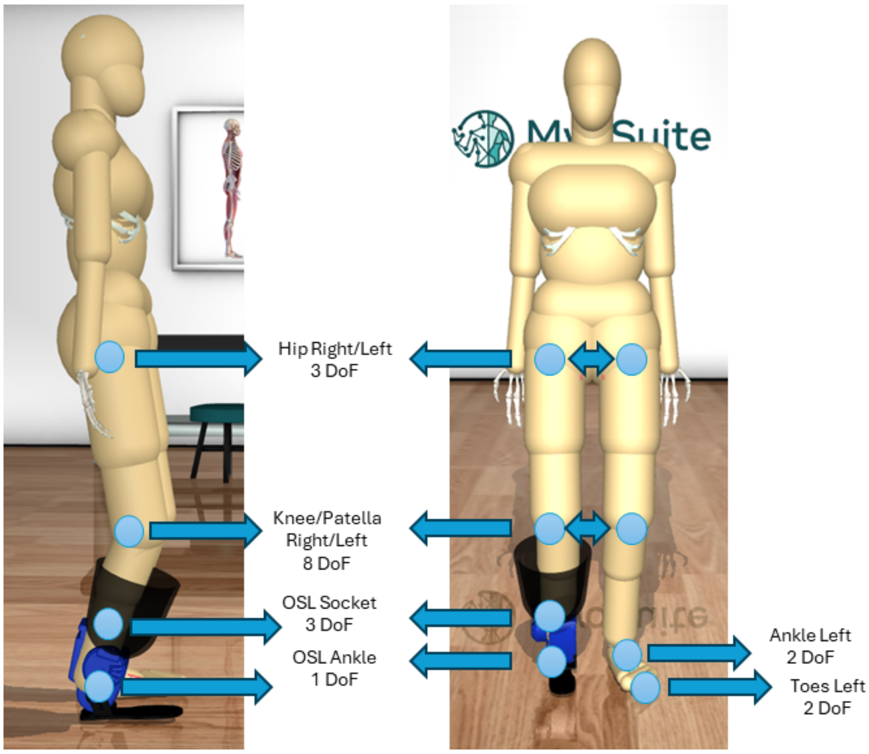

Figure 2. Lower-Limb MS model with number of DoF.

the virtualised model, the physical caracteristics of the motors and range of movement were gathered and can be seen in detail in Table 2.

Table 3 provides an overview of the joints included in the model, their names, the number to be accessed by the MyoSuite simulation environment, and corresponding reference images. Additionally, the range of motion for each joint, essential for simulating realistic movement, is detailed in Table 4. These tables collectively highlight the comprehensive nature of the MS model, designed to replicate mechanics of human walkin locomotion.

By integrating the MS model with the MyiSuite interface, is possible to develop and test control strategies for prosthetic devices. This integration ensures that the model can represents human anatomy and movement, while providing a versatile platform for enhancing the design and functionality of prosthetics. With the use of control algorithms with machine learning, it is possible to improve the adaptability and performance of prosthetic limbs.

3.2. Reinforcement Learning Approach

The algorithm used for locomotion pattern generation was DEP-RL, by integrating Differential Extrinsic Plasticity *DEP* with reinforcement learning *RL* to enhance learning efficiency and performance with complex multivariable systems, in this case, musculoskeletal systems. DEP is a mechanism inspired by Hebbian learning principles, adjusting interaction between sensors and actuators based on their mutual activity [44].DEP enhances the Hebbian learning by incorporating environmental feedback into the learning process, allowing the system to adapt the behaviour based on the outcome of the selected actions. The control matrix C is based on the correlations between state changes and actions over time, give by equation 1, where τ controls the time scale and $\dot{s}_t$ and $\dot{s}_{t-\delta t}$ are the state velocities at times t and $t - \delta t$.

$$ {}_{\tau}\dot{C} = \dot{s}_t \dot{s}_{t-\delta t}^T - C \qquad (1) $$

DEP method is integrated with a RL algorithm, in this case, MPO (Maximum a Posteriori Policy Optimization) [45]. With the combination, DEP method is used for exploration

Table 2. Technical specifications of the OSL actuator, range of movement and MyoSuite integration [41,42]

Specification	**Ankle Motor**
Brand	Dephy Modified T-motor U8-KV100
Gear Ratio	58.4:1
Peak Torque (Nm)	4.2 Nm (motor) 245.3 Nm (with pair)
Continuous Torque (Nm)	0.14 Nm/A (motor) 8.176 Nm/A (with Pair)
Range of Movement (degrees)	30° (-15° to 15°)
Energetic Usage	Max Continuous Current: 7.7 A Peak Current (20 s): 28.8 A, Fused at 30 A Voltage: 36 V
MyoSuite Details	Gear: 58.4 Control Range: -2.88 to 2.88

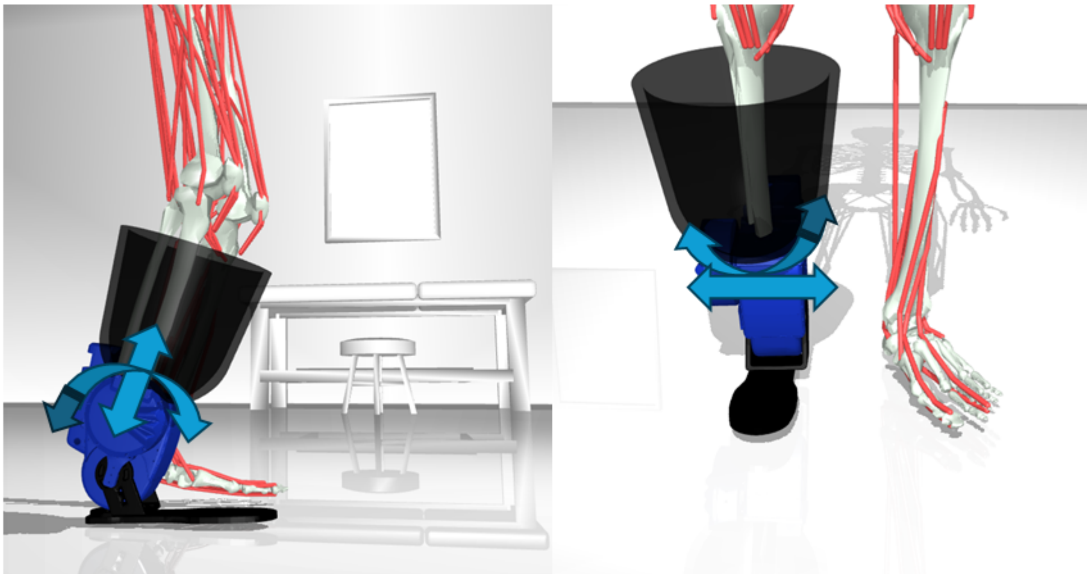

Figure 3. Socket DoF based on the dynamics of a generic socket movement [43].

Table 3. Joints on the model and range of movement. Joints are mirrored between left and right foot, except by Ankle prosthetic devices, which performs ankle extension and flexion.

Name	Degrees of Movement
Hip Extension/Flexion	-10 ° to 120°
Hip adduction/abduction	-30° to 60°
Hip Rotation (internal to external)	-45° to 60°
Knee Flexion	0 to 135°
Knee Rotation (internal to external)	-20° to 10°
Ankle Extension/Flexion	-50° to 30°
Ankle Inversion/Eversion	-30° to 60°
Toes Extension/Flexion	-70° to 45°
Transtibial prosthetic Motor	-50° to 30°

Table 4. Details of the DoF in the Model, for details of the movement, referred to the mentioned Figures

Joint	Name	Details
0	Root	Reference point
1 to 6	Translation in the world	Movement in the world
7	Hip Flexion Right	Figure 2
8	Hip Adduction Right	Figure 2
9	Hip Rotation Right	Figure 2
10	Tibia angle translation 2	Figure 4
11	tibia angle translation 1	Figure 4
12	Knee angle flexion/extension	Figure 2
13	Tibia angle rotation 2	Figure 4
14	Tibia angle rotation 1	Figure 4
15	Socket Piston	Figure 3
16	Socket Rotation 1	Figure 3
17	Socket Rotation 2	Figure 3
18	Socket Rotation 3	Figure 3
19	OSL Ankle Angle Right	Figure 2
20	Patella Angle Right Beta Translation 2	Figure 4
21	Patella Angle Right Beta Translation 1	Figure 4
22	Patella Angle Right Beta Rotation 1	Figure 4
23	Hip Flexion Left	Figure 2
24	Hip Adduction Left	Figure 2
25	Hip Rotation Left	Figure 2
26	Knee Angle Left Translation 2	Figure 4
27	Knee Angle Left Translation 1	Figure 4
28	Knee Angle Left	Figure 4
29	Knee Angle Left Rotation 2	Figure 4
30	Knee Angle Left Rotation 3	Figure 4
31	Ankle Angle Left	Figure 2
32	Subtalar Angle Left	Figure 2
33	MTP Angle Left	Figure 2
34	Patella Angle Left Beta Translation 2	Figure 4
35	Patella Angle Left Beta Translation 1	Figure 4
36	Patella Angle Left Beta Rotation 1	Figure 4

Table 5. Muscles in the simulated model with their primary function and joint movement

Real Name	Primary Function	Joint Movement	Removed on Right
Adductor brevis	Thigh Adduction	Hip	No
Adductor longus	Thigh Adduction, rotation	Hip	No
Adductor magnus distal	Pelvis stabilizer	Hip, Knee	No
Adductor magnus ischial	Pelvis stabilizer	Hip, Knee	No
Adductor magnus mid	Pelvis stabilizer	Hip, Knee	No
Adductor magnus proximal	Pelvis stabilizer	Hip, Knee	No
Bicep Femoral Long Head	Thigh extension, rotation, Knee flexion	Hip, Knee	No
Bicep Femoral Short Head	Knee Rotation	Knee	No
Extensor Digitorum Longus	Digits movement	Digits	Yes
Extensor Hallucis Longus	Digits movement	Digits	Yes
Flexor Digitorum Longus	Digits movement	Digits	Yes
Flexor Hallucis Longus	Digits movement	Digits	Yes
Gastrocnemious Lateral	Plantar Flexion, knee Flexion	Ankle	Yes
Gastrocnemious medial	Plantar Flexion, knee Flexion	Ankle	Yes
Gluteus Maximus 1	Thigh Rotation	Hip	No
Gluteus Maximus 2	Thigh Rotation	Hip	No
Gluteus Maximus 3	Thigh Rotation	Hip	No
Gluteus Medius 1	Hip Abduction	Hip	No
Gluteus Medius 2	Hip Abduction	Hip	No
Gluteus Medius 3	Hip Abduction	Hip	No
Gluteus Minimus 1	Hip Abduction and Stabilizer	Hip	No
Gluteus Minimus 2	Hip Abduction and Stabilizer	Hip	No
Gluteus Minimus 3	Hip Abduction and Stabilizer	Hip	No
Gracilis Muscle	Thigh abduction, knee flexion	Hip, Knee	No
Iliacus Muscle	Femur Rotation	Hip	No
Peroneus Brevis	Foot Eversion	Ankle	Yes
Peroneus Long	Foot Eversion	Ankle	Yes
Piriformis	Thigh Rotation	Hip	No
Psoas Iliaco	Hip Rotation	Hip	No
Rectus Femoris	Knee Extension	Knee	No
Sartorius Muscle	Hip and Knee Movement	Hip, Knee	No
Semimembranosus	Hip and Knee Movement	Hip, Knee	No
Semitendinosus	Hip and Knee Movement	Hip, Knee	No
Soleus Muscle	Plantar Flexion	Ankle	Yes
Tensor Fasciae Latae	Knee Rotation	Knee	No
tibialis anterior	Plantar Dorsiflexion	Ankle	Yes
tibialis posterior	Plantar Flexion and inversion	Ankle	Yes
vastus intermedius	Knee Extension	Knee	No
vastus lateralis	Knee Extension	Knee	No
vastus medialis	Knee Extension	Knee	No
Ankle Motor	Ankle Flexion	Ankle	

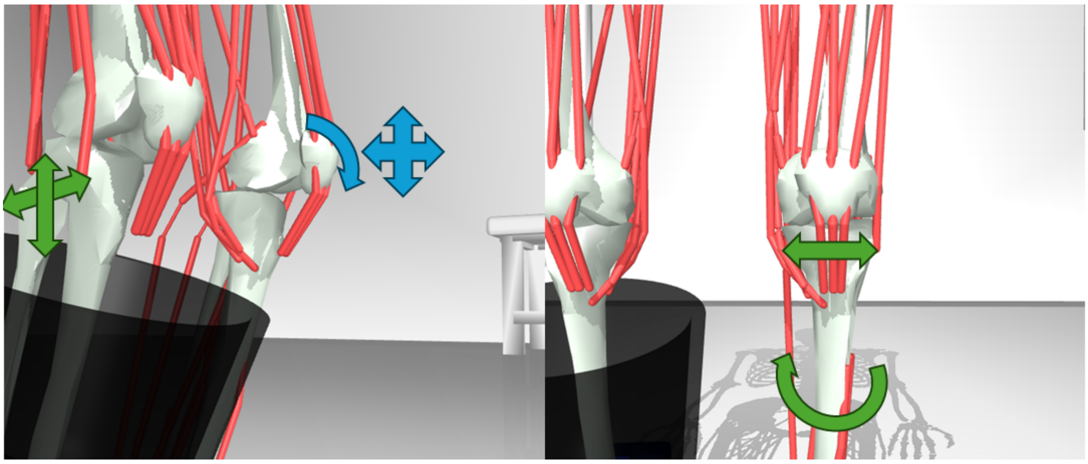

Figure 4. Detailed DoF in the Patella (blue marked) and Tibia (green marked). Displacement passively performed by leg muscles.

noise, improving learning efficiency in high-dimensional action spaces, like the current model, involving muscle activations. DEP-RL alternates between the RL policy π and DEP driven exploration in each episode, based on a stochastic switching probability p_{switch} [46]. The strategy enables the RL algorithm to use DEP efficient exploration of the action space, and force it to go out of specific limits[47], enhancing the learning process for locomotion and accelerating the convergence to robust strategies.

After starting the control matrix C in zeros, updates are merely based on the environment's response, using connection to the future state to predict and adjust actions. By mapping the sensors directly to the actuatores, new actions are generated, using a tanh function, commonly used to ensure the actions remain in a bounded range, as seen in equation 2, where k is an amplification constant, and h_t is a time-dependent bias term.

$$a_t = tanh(kCs_t + h_t) \tag{2}$$

DEP-RL application to locomotion tasks has shown higher accuracty results compared to other RL methods on locomotion tasks in sample efficiency and adaptability, with no previous experimental data.

The training process involves iteratively updating the policy based on the rewards obtained from interactions with the simulation environment, optimizing the agent's ability to control the prosthetic limb effectively.

The observation space in this study includes different variables describing the current status of the MS model in the world. These variables include joint positions, velocities, muscle activations, and sensory inputs from the prosthetic limb. By providing a detailed representation of the model's state, the observation space allows the RL agent to make informed decisions based on the current biomechanical configuration.

The action space consists of the control signals sent to the muscles and the prosthetic actuator. This high-dimensional action space enables the agent to finely tune the muscle forces and actuator inputs, facilitating precise and adaptive control of the lower-limb during locomotion.

The reward function guides the agent toward the desirable behaviour by providing feedback on its performance. This was designed to promote stable walking by incorporating the following elements:

- **Forward Progression:** Incentivise a constant target velocity during locomotion. As seen in Eq. 3, reward is calculated on the difference between current velocity V_y, V_x and target velocities $targetV_y$, $targetV_x$. To keep a straight line, $targetV_x$ remains 0.

$$R_{vel} = \exp\left(-(targetV_y - v_y)^2\right) + \exp\left(-targetV_x - v_x)^2\right) \tag{3}$$

- **Cyclic Hip:** Encourage the agent to perform a cyclic motion pattern in the hip joints, characteristic of natural walking gait. To calculate the pattern, the phase ($phase\_var$)

of walking must be calculated, ensuring that the movement of the two hips are out of phase. This is calculated in Eq. 4, where it takes the current step count (*steps*) and the hip period (*hip_period*) for each leg. This value is used in Eq. 5 to calculate the desired angles for each hip. Using a coefficient 0.8 in the cosine function [36] ($\approx 45°$), the agent is rewarded at moving out of phase. The negative reward, as detailed in Eq. 6 decreases as the difference between the desired angles (*des_angles*) and actual angles increases, penalised for deviating from the desired cyclic pattern. Smaller differences results in a less negative reward.

$$\text{phase\_var} = \left(\frac{\text{steps}}{\text{hip\_period}}\right) \mod 1 \tag{4}$$

$$\text{des\_angles} = \begin{pmatrix} 0.8\cos(2\pi \cdot \text{phase\_var} + \pi) \\ 0.8\cos(2\pi \cdot \text{phase\_var}) \end{pmatrix} \tag{5}$$

$$R_{ch} = \left\| \text{des\_angles} - \begin{pmatrix} \theta_{\text{hip\_flexion\_l}} \\ \theta_{\text{hip\_flexion\_r}} \end{pmatrix} \right\| \tag{6}$$

- **Reference Rotation:** Eq. 7 incentivise the agent to maintain a current stable orientation (*curr_pos*) close to the target rotational position (*target_rot*), which is selected to be the initial position.

$$R_{rr} = \exp\left(-\left\|5.0 \cdot (curr\_pos - \text{target\_rot})\right\|\right) \tag{7}$$

- **Symmetry:** Encourages symmetrical movement between the prosthetic and organic limbs. Eq. 8 measures the differece in movement patterns between the left and right legs by using the mean of the GRF sensors data after a whole step cycle, providing a less negative reward, if the difference is minimum, ensuring the same amount of steps have been taken.

$$R_{sym} = exp(-\left|\text{left\_leg\_movement} - \text{right\_leg\_movement}\right|) \tag{8}$$

- **Step Length:** Reward an appropiate length of each step, by ensuring both legs produced a balanced and consistent distance, between each other. Calculation of each leg step distance are performed in Eq. 9, where it is measured by the absolute difference between the position of the $hip\_flexion$ and $knee\_angle$ joints. By calculating the difference between both legs step length ($step\_length\_diff$), both legs will produce similar movements. The reward will be closer to 1, the less difference exists between the movement of both legs (10).

$$\text{step\_length} = |q_{\text{hip\_flexion}} - q_{\text{knee\_angle}}| \tag{9}$$

$$R_{sl} = \exp\left(-\text{step\_length\_diff}\right) \tag{10}$$

- **Joint Angle Reward:** Encourages the agent to maintain appropiate joint angles for specific joints. Eq. 11 calculate the reward based on the mean absolute of the left and right hip adduction and rotation joint angles (θ_joint), to avoid high deviation from their neutral positions.

$$R_{ja} = \exp\left(-5 \cdot \text{mean}\left(|\theta_{joint}|\right)\right) \tag{11}$$

- **Activation Magnitude:** Penalise excessive muscle activations, promoting energy-efficient movements. Eq. 12 calculate the reward as the norm of the activation signals (amount of muscles activating) divided by the number of actuators, ensuring the agent uses the minimum necessary muscle activations for the locomotion action.

$$\text{act\_mag} = \begin{cases} \frac{\|\mathbf{act}\|}{number\_actuators} & \text{if } number\_actuators \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

At the end each of the episode a small penalization was given to the agent, summing if the agent falled down, it produced a high negative penalization, ending the episode too. Giving the reward function seen in Eq. 13, multiplied with twekable weights by the user.

$$R = 5.0R_{vel} - 10.0R_{ch} + 10.0R_{rr} + 10.0R_{sym} + 5.0R_{ja} + 5.0R_{sl} - 100.0R_{fall} - 20.0R_{done} + R_{act\_mag} \quad (13)$$

3.3. Testing and Evaluation

To ensure effectiveness and robustness of the prosthetic control system, a series of tests were conducted. These tests aimed to evaluate the system's balance, learning performance, energy efficiency, and adaptabilit to various terrains.

(1) **Balance Test Objective:** Train the model to maintain a static upright position and measure its stability, by calculating the centre of mass (CoM) deviation. A second test was perfomed, with a disturbance in the environment, produced by adding an external force at the pelvis body at a random horizontal direction, at the 6 seconds mark to verify models' reaction.
(2) **Reinforcement Learning Training and Performance Objective:** Track the learning curve and convergence of the RL algorithm and evaluate model's performance in walking tasks post-training. Evaluation will be based on achieved velocity and distance covered in 10 seconds.
(3) **Adaptability and Robustness Testing Objective:** Evaluate the agent performance on different terrains, including flat surfaces, uneven surfaces, and ramps. The evaluation is based on distance travelled without falling and achieved velocity.

4. Results

4.1. Balance

The model was first trained to keep a standing position across time, capable of enduring disturbances. To keep track of the stability of the model, the CoM was tracked during 10 seconds. Model standing position can be seen on Figure 5.c, where the muscles in the body were activated to keep both feet on the ground while keeping the CoM in the same position. Figure 5.a shows how even when the body keeps the model in the upright position, it leaves the error close to 2% which could be improved minimizing the treshold for the correct position. Contrasting, Figure 5.b shows how at the mark of 6 seconds, the disturbance seems achieving a higher displacement of the CoM, however, this movement, causes the model to recover the original position.

4.2. Learning Performance

To understand the RL algorithm capacity to solve the described problems, the score vs Steps or Epochs was tracked. First balance problem was solved in $2e^8$ epochs, achieving a cumulative score of 100,000. This score was the cumulative reward of a portion of $2e^5$ steps, as seen in Figure 6.a.

Second Test was to achieve walking motion on even floor, to later adapt it to rough terrain and terrain with hills on the way. Walking motion training performance can be seen in Figure 6.b, with some downhills in the curve at the score, starting at epoch $2e^8$. These down curves are obtained due to modifications in the reward function to improve walking locomotion either on the even floor and other terrains. Since at modifying the terrain, the usual walking locomotion needed some tweaks, to adapt. This adaptation caused by the RL exploitation step, caused less higher rewards, before achieving optimal values.

First walking motion was performed on even floor with no disturbances. To measure the performance, velocity and distance covered in straight line (y direction in the plane) were measured. After $2e^8$ epochs, the training results were optimal in achieving a walking locomotion with a high capability of achieving the desired velocity of 1.2 m/s as seen in Figure 7.a. Deviation of the velocity was of 1.2%, measuring last 200 steps. First steps were not taken into

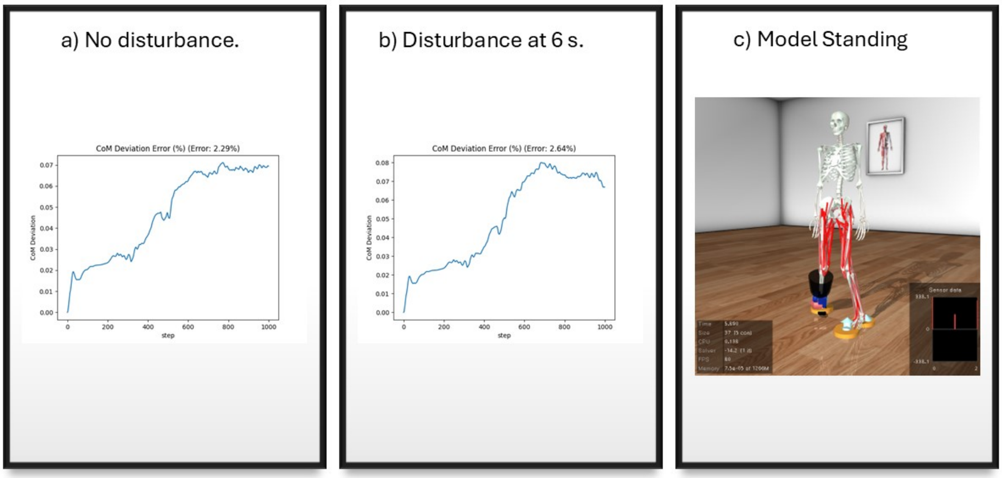

Figure 5. Results of Standing Trials. Showing how the model learned to stay still even with a disturbance applied on the pelvis. Disturbance created in the environment by adding a external force at a random horizontal direction at the mark of 6 seconds.

Figure 6. Epochs trained to get optimal results. Higher score means, better performance, according to the reward function.

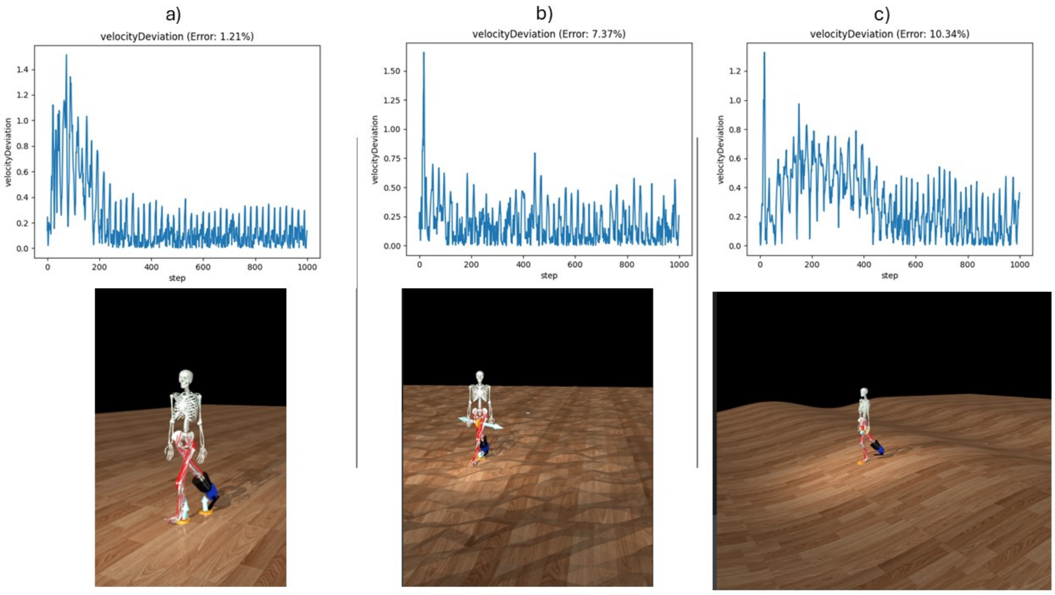

Figure 7. Walking motion and velocity achieved during each of the test scenarios. a)Even Floor, b) Uneven Floor, c) Hills showing.

account due to the models' requirement to accelerate until desired speed. Distance covered was a always a little bit more than 10 meters, as expected due to the speed, starting from a resting point.

4.3. Adaptability

Model was trained and evaluated for motion performance under rough and hilly terrain. Model easily performed motion under the rough terrain stage seen in Figure 7.b, designed with negative and positive elevations of -12 cm and $2cm$. Achieved travelled distance over 10 seconds was 10.8 meters with a velocity deviation error of 7.37%.

Hill Terrain Stage presented a higher difficulty for the model, requiring further training in the neural network. Same velocity was required to achieve the motion under uphills and downhills as seen in Figure 7.c, modifying the speed to keep the model with stable walking motion. As seen in the velocity deviation graphic, the deviation is higher at some points due to this variation on the elevation of the terrain, increasing the error up to 10.34 %.

5. Discussion

The study underscores the importance of integrating detailed MS models with RL to advance prosthetic control systems. The results highlight the current model's ability to achieve basic walking motion but reveal areas needing further improvement. The model demonstrated adequate standing motion and center of mass (CoM) stability during disturbances, suggesting robustness in maintaining balance. However, the leg movements did not achieve the desired outcomes. Although the cyclic motion was consistent, the prosthetic leg remained positioned behind for stability, while the organic leg led in movement and direction, reflecting an imbalance in functionality. This outcome is a common problem seen in different DRL applications, where the agent learns how to perform outstanding, getting high rewards, with unexpected movements, which indicates a need to refine the control system to improve leg coordination and movement symmetry. Future work should focus on enhancing the model's control system to replicate more natural human-like gait patterns by adjusting the control algorithms to perform cyclic and constant motion, while validating the model with real-world data. Enhanced stability and coordination will make the prosthetic limb more effective in diverse real-life situations.

6. Conclusion

Our study successfully implements a reinforcement learning-based control system for a virtual musculoskeletal model with a transtibial prosthesis, consisting of 70 muscles and one electrical actuator at the ankle, achieving forward locomotion without falling, although it has yet to resemble natural human gait. This paper highlights our methodology and demonstrates the potential of integrating RL with a detailed musculoskeletal model to handle high-dimensional control tasks in prosthetics. The model was able to achieve standing motion, even with a external force pushing the model as a disturbance, achieving a CoM error of only 2% after from the initial point. By testing the walking locomotion in different environments, such as hill, uneven and plain terrain, it showed the model's capability to be resilient against floor disturbances, with problems only at achieving straight line motion, with a deviation of 10%, however, with no falling events, and achieving a velocity accuraty of 92.67%. The findings pave the way for further refinement of the control systems, aiming to achieve more realistic and efficient locomotion with a prosthetic ankle.

Disclosure statement

The authors certify that, to the best of our knowledge, there are no competing interests to declare.

Funding

The authors would like to thank Tecnologico de Monterrey for the financial support provided through the 'Challenge-Based Research Funding Program 2023', Project ID #IJXT070-23EG99001, titled 'Complex Thinking Education for All (CTE4A): A Digital Hub and School for Lifelong Learners.'

Note on Collaborators

Adan Domínguez-Ruiz. MsC in Robotics by the University of Bristol, currently studying the Ph.D in Engineering Sciences from the Instituto Tecnológico de Monterrey. Robotics and Digital Systems Engineer by the same university in 2016. His lines of research are in bionics, reinforcement learning, cognitive systems, human-robotic interaction, Education 4.0 and biomedical sciences.

Oscar Loyola holds a Ph.D. in Engineering Sciences with a specialization in Automation from the University of Santiago, Chile (2021). He obtained a Master's degree in Engineering Sciences with a specialization in Electrical Engineering from the same university in 2017, and a Master's degree in Biomedical Engineering from the International University of Valencia in 2022. He is a Civil Electronic Engineer from the Ibero-American University of Science and Technology and also holds a degree as an Automation Engineer in Industrial Processes. Currently, he is part of the Robotics and Industry 4.0 research group. His areas of work include reinforcement learning in robots, cognitive systems, control theory, and systems modeling.

Edgar-Omar López Caudana. Ph.D. in Communications and Electronics from the Instituto Politécnico Nacional. Research Professor of the Reasoning for Complexity Group of the Institute for the Future of Education. His main lines of research are in Higher Education and Social Robotics, a line which he has largely developed in recent years with a focus on health and education (from preschool to postgraduate levels). He focuses his projects on the implementation and design of educational initiatives, with innovation in Education 4.0, as a means of transformation and social impact. He has presented several research articles, collaborated in book chapters, and participated in different conferences.

Pedro Ponce-Cruz.Professor and researcher at Tecnológico de Monterrey, Mexico. Ph.D. in Engineering with expertise in control systems, artificial intelligence, and machine learning. His research areas encompass fuzzy logic, neural networks, renewable energy, and the application of advanced computational methods for engineering problem-solving. He is widely recognized for his contributions to the development of intelligent control systems and his work has been cited extensively in academic literature. Additionally, he has authored several books and articles that are utilized by researchers and practitioners worldwide in the fields of control theory, automation, and robotics.

7. Data availability statement

Data can be found in the following GitHub Repository, only with the last NN weights working. DigitalTwin Repository

8. References

References

[1] Domínguez-Ruiz A, López-Caudana EO, Lugo-González E, et al. Low limb prostheses and complex human prosthetic interaction: A systematic literature review. Frontiers in Robotics and AI. 2023;10. Available from: `https://www.frontiersin.org/articles/10.3389/frobt.2023.1032748`.

[2] George-Reyes CE, López-Caudana EO, Ramírez-Montoya MS, et al. Pensamiento computacional basado en realidad virtual y razonamiento complejo: caso de estudio secuencial. Revista de Educación a Distancia (RED). 2023 Jan;23(73). Number: 73; Available from: `https://revistas.um.es/red/article/view/540841`.

[3] George-Reyes CE, Peláez Sánchez IC, Glasserman-Morales LD, et al. The Metaverse and complex thinking: opportunities, experiences, and future lines of research. Frontiers in Education. 2023 May;8. Publisher: Frontiers; Available from: `https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2023.1166999/full`.

[4] Delp S, Loan J, Hoy M, et al. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. IEEE Transactions on Biomedical Engineering. 1990 Aug;37(8):757–767. Conference Name: IEEE Transactions on Biomedical Engineering; Available from: `https://ieeexplore.ieee.org/document/102791`.

[5] Seth A, Sherman M, Reinbolt JA, et al. OpenSim: a musculoskeletal modeling and simulation framework for in silico investigations and exchange. Procedia IUTAM. 2011;2:212–232. Available from: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4397580/`.

[6] Anderson FC, Pandy MG. Dynamic optimization of human walking. Journal of Biomechanical Engineering. 2001 Oct;123(5):381–390.

[7] Fernandez JW, Mithraratne P, Thrupp SF, et al. Anatomically based geometric modelling of the musculo-skeletal system and other organs. Biomechanics and Modeling in Mechanobiology. 2004 Mar;2(3):139–155.

[8] Rosen R, von Wichert G, Lo G, et al. About The Importance of Autonomy and Digital Twins for the Future of Manufacturing. IFAC-PapersOnLine. 2015 Jan; 48(3):567–572. Available from: `https://www.sciencedirect.com/science/article/pii/S2405896315003808`.

[9] Sartori M, Farina D, Lloyd DG. Hybrid neuromusculoskeletal modeling to best track joint

moments using a balance between muscle excitations derived from electromyograms and optimization. Journal of Biomechanics. 2014 Nov;47(15):3613–3621.

[10] Mishra UA. Learning Control Policies for Imitating Human Gaits. ArXiv. 2021 May;Available from: https://www.semanticscholar.org/paper/043be377aba464a5aa05d49bcd3028e1dcedb1ba.

[11] Kaymak C, Ucar A, Guzelis C. Development of a New Robust Stable Walking Algorithm for a Humanoid Robot Using Deep Reinforcement Learning with Multi-Sensor Data Fusion. Electronics. 2023 Jan;12(3):568. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute; Available from: https://www.mdpi.com/2079-9292/12/3/568.

[12] Singh RP, Benallegue M, Morisawa M, et al. Learning Bipedal Walking On Planned Footsteps For Humanoid Robots. In: 2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids); Nov.; 2022. p. 686–693. ISSN: 2164-0580.

[13] Bartlett HL, King ST, Goldfarb M, et al. A Semi-Powered Ankle Prosthesis and Unified Controller for Level and Sloped Walking. IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society. 2021;29:320–329.

[14] Tucker MR, Olivier J, Pagel A, et al. Control strategies for active lower extremity prosthetics and orthotics: a review. Journal of NeuroEngineering and Rehabilitation. 2015 Dec;12(1):1–30. Number: 1 Publisher: BioMed Central; Available from: https://jneuroengrehab.biomedcentral.com/articles/10.1186/1743-0003-12-1.

[15] Hicks JL, Uchida TK, Seth A, et al. Is my model good enough? Best practices for verification and validation of musculoskeletal models and simulations of movement. Journal of Biomechanical Engineering. 2015 Feb;137(2):020905.

[16] Seth A, Hicks JL, Uchida TK, et al. OpenSim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. PLoS Computational Biology. 2018 Jul;14(7):e1006223. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6061994/.

[17] Whitmore MW, Hargrove LJ, Perreault EJ. Gait Characteristics When Walking on Different Slippery Walkways. IEEE transactions on bio-medical engineering. 2016 Jan; 63(1):228–239.

[18] Altinkaynak ES, Braun DJ. A Phase-Invariant Linear Torque-Angle-Velocity Relation Hidden in Human Walking Data. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2019 Apr;27(4):702–711. Conference Name: IEEE Transactions on Neural Systems and Rehabilitation Engineering.

[19] Kim M, Chen T, Chen T, et al. An Ankle–Foot Prosthesis Emulator With Control of Plantarflexion and Inversion–Eversion Torque. IEEE Transactions on Robotics. 2018 Oct; 34(5):1183–1194. Available from: https://ieeexplore.ieee.org/document/8372941/.

[20] Leestma JK, Golyski PR, Smith CR, et al. Linking whole-body angular momentum and step placement during perturbed human walking. The Journal of Experimental Biology. 2023 Mar;226(6):jeb244760. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10112983/.

[21] Rajagopal A, Dembia CL, DeMers MS, et al. Full-Body Musculoskeletal Model for Muscle-Driven Simulation of Human Gait. IEEE Transactions on Biomedical Engineering. 2016 Oct;63(10):2068–2079. Available from: http://ieeexplore.ieee.org/document/7505900/.

[22] Hodossy BK, Farina D. Shared Autonomy Locomotion Synthesis with a Virtual Powered Prosthetic Ankle. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2023;:1–1Conference Name: IEEE Transactions on Neural Systems and Rehabilitation Engineering; Available from: https://ieeexplore.ieee.org/document/10332161.

[23] Carloni R, Luinge R, Raveendranathan V. The gait1415+2 OpenSim musculoskeletal model of transfemoral amputees with a generic bone-anchored prosthesis. Medical Engineering & Physics. 2024 Jan;123:104091. Available from: https://www.sciencedirect.com/science/article/pii/S1350453323001467.

[24] Wang H, Caggiano V, Durandau G, et al. MyoSim: Fast and physiologically realistic

MuJoCo models for musculoskeletal and exoskeletal studies. In: 2022 International Conference on Robotics and Automation (ICRA); May; 2022. p. 8104–8111. Available from: https://ieeexplore.ieee.org/abstract/document/9811684.

[25] Pickle NT, Grabowski AM, Jeffers JR, et al. The Functional Roles of Muscles, Passive Prostheses, and Powered Prostheses During Sloped Walking in People With a Transtibial Amputation. Journal of Biomechanical Engineering. 2017 Nov;139(11):1110051–11100511.

[26] Silverman AK, Neptune RR. Muscle and prosthesis contributions to amputee walking mechanics: a modeling study. Journal of Biomechanics. 2012 Aug;45(13):2271–2278.

[27] Geijtenbeek T. The Hyfydy Simulation Software ; 2021. Available from: https://hyfydy.com.

[28] Todorov E, Erez T, Tassa Y. MuJoCo: A physics engine for model-based control. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems; Oct.; Vilamoura-Algarve, Portugal. IEEE; 2012. p. 5026–5033. Available from: http://ieeexplore.ieee.org/document/6386109/.

[29] Geijtenbeek T. SCONE: Open Source Software for Predictive Simulation of Biological Motion. Journal of Open Source Software. 2019 Jun;4(38):1421. Available from: https://joss.theoj.org/papers/10.21105/joss.01421.

[30] Caggiano V, Wang H, Durandau G, et al. MyoSuite – A contact-rich simulation suite for musculoskeletal motor control ; 2022. ArXiv:2205.13600 [cs]; Available from: http://arxiv.org/abs/2205.13600.

[31] Rattanasak A, Uthansakul P, Uthansakul M, et al. Real-Time Gait Phase Detection Using Wearable Sensors for Transtibial Prosthesis Based on a kNN Algorithm. Sensors. 2022 Jan;22(11):4242. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute; Available from: https://www.mdpi.com/1424-8220/22/11/4242.

[32] Kim M, Hargrove LJ. A gait phase prediction model trained on benchmark datasets for evaluating a controller for prosthetic legs. Frontiers in Neurorobotics. 2023;16. Available from: https://www.frontiersin.org/articles/10.3389/fnbot.2022.1064313.

[33] Shafiul Hasan SM, Siddiquee MR, Atri R, et al. Prediction of gait intention from pre-movement EEG signals: a feasibility study. Journal of NeuroEngineering and Rehabilitation. 2020 Apr;17(1):50. Available from: https://doi.org/10.1186/s12984-020-00675-5.

[34] Mohammedalamen M, Khamies WD, Rosman B. Transfer Learning for Prosthetics Using Imitation Learning ; 2019. ArXiv:1901.04772 [cs]; Available from: http://arxiv.org/abs/1901.04772.

[35] Berg C, Caggiano V, Kumar V. SAR: Generalization of Physiological Agility and Dexterity via Synergistic Action Representation ; 2023. ArXiv:2307.03716 [cs]; Available from: http://arxiv.org/abs/2307.03716.

[36] Schumacher P, Haeufle DFB, Büchler D, et al. DEP-RL: EMBODIED EXPLORATION FOR REINFORCEMENT LEARNING IN OVERACTUATED AND MUSCULOSKELETAL SYSTEMS; 2023.

[37] Ficanha E, Aramizo Ribeiro G, Knop L, et al. Estimation of the 2-DOF Time-Varying Impedance of the Human Ankle. In: 2017 Design of Medical Devices Conference; Apr.; Minneapolis, Minnesota, USA. American Society of Mechanical Engineers; 2017. p. V001T05A001.

[38] Chumacero E, Al Masud A, Isik D, et al. Advances in Powered Ankle-Foot Prostheses. Critical Reviews in Biomedical Engineering. 2018;46(3):185–200. Available from: http://www.dl.begellhouse.com/journals/4b27cbfc562e21b8,3ce1ac5b6e88c9ed,4619161b05664487.html.

[39] Surgeons AAoO. Atlas of Limb Prosthetics: Surgical, Prosthetic, and Rehabilitation Principles. Mosby Year Book; 1992. Google-Books-ID: xPtsAAAAMAAJ; Available from: https://www.oandplibrary.org/alp/.

[40] Raveendranathan V, Kooiman VGM, Carloni R. Musculoskeletal model of osseointegrated transfemoral amputees in OpenSim. PLOS ONE. 2023 Sep;18(9):e0288864. Publisher: Public Library of Science; Available from: https://journals.plos.org/

plosone/article?id=10.1371/journal.pone.0288864.
[41] Azocar AF, Mooney LM, Duval JF, et al. Design and clinical implementation of an open-source bionic leg. Nature Biomedical Engineering. 2020 Oct;4(10):941–953. Available from: https://www.nature.com/articles/s41551-020-00619-3.
[42] Lee UH, Pan CW, Rouse EJ. Empirical Characterization of a High-performance Exterior-rotor Type Brushless DC Motor and Drive. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); Nov.; Macau, China. IEEE; 2019. p. 8018–8025. Available from: https://ieeexplore.ieee.org/document/8967626/.
[43] McGeehan MA, Adamczyk PG, Nichols KM, et al. A simulation-based analysis of the effects of variable prosthesis stiffness on interface dynamics between the prosthetic socket and residual limb. Journal of Rehabilitation and Assistive Technologies Engineering. 2022 Jul;9:20556683221111986. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9289901/.
[44] Der R, Martius G. Novel plasticity rule can explain the development of sensorimotor intelligence. Proceedings of the National Academy of Sciences of the United States of America. 2015 Nov;112(45):E6224–E6232. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4653169/.
[45] Abdolmaleki A, Springenberg JT, Tassa Y, et al. Maximum a Posteriori Policy Optimisation ; 2018. ArXiv:1806.06920 [cs, math, stat]; Available from: http://arxiv.org/abs/1806.06920.
[46] Pislar M, Szepesvari D, Ostrovski G, et al. When should agents explore? Oct.; 2021. Available from: https://openreview.net/forum?id=dEwfxt14bca.
[47] Martius G, Hostettler R, Knoll A, et al. Compliant control for soft robots: Emergent behavior of a tendon driven anthropomorphic arm. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); Oct.; 2016. p. 767–773. ISSN: 2153-0866; Available from: https://ieeexplore.ieee.org/document/7759138.