Energy Regeneration and Environment Sensing for Robotic Leg Prostheses and Exoskeletons

by
Brokoslaw Laschowski

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2021

EXAMINING COMMITTEE MEMBERSHIP

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr. Max Donelan Department of Biomedical Physiology and Kinesiology, Simon Fraser University
Internal/External Member	Dr. Arash Arami Department of Mechanical and Mechatronics Engineering, University of Waterloo
Internal Member 1	Dr. Ning Jiang Department of Systems Design Engineering, University of Waterloo
Internal Member 2	Dr. James Tung Department of Mechanical and Mechatronics Engineering, University of Waterloo
Supervisor	Dr. John McPhee Department of Systems Design Engineering, University of Waterloo

AUTHOR'S DECLARATION

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

STATEMENT OF CONTRIBUTIONS

This thesis contains some material taken from multi-author papers, for which I was lead author on all original research articles, including:

1. Maryniak A, Laschowski B, and Andrysek J. (2018). Technical overview of osseointegrated transfemoral prostheses: Orthopaedic surgery and implant design centered. ASME Journal of Engineering and Science in Medical Diagnostics and Therapy, 1(2), pp. 020801-020801-7. DOI: 10.1115/1.4039105.

Contribution of B. Laschowski: assisted with the review of osseointegrated knee prostheses.

2. Laschowski B and Andrysek J. (2018). Electromechanical design of robotic transfemoral prostheses. ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC-CIE). DOI: 10.1115/DETC2018-85234.

Contribution of J. Andrysek: assisted with the device design and performance review.

3. Laschowski B, McPhee J, and Andrysek J. (2019). Lower-limb prostheses and exoskeletons with energy regeneration: Mechatronic design and optimization review. ASME Journal of Mechanisms and Robotics, 11(4), pp. 040801. DOI: 10.1115/1.4043460.

Contribution of J. Andrysek: assisted with the device design and performance review.

4. Laschowski B, McNally W, Wong A, and McPhee J. (2019). Preliminary design of an environment recognition system for controlling robotic lower-limb prostheses and exoskeletons. IEEE International Conference on Rehabilitation Robotics (ICORR), pp. 868-873. DOI: 10.1109/ICORR.2019.8779540.

Contribution of W. McNally: technical support the neural network design and training.

5. Laschowski B, Razavian RS, and McPhee J. (2019). Modelling and biomechanical evaluation of stand-to-sit motions: Implications for energy-efficient design of prostheses and exoskeletons. International Society of Biomechanics Congress.

Contribution of RS. Razavian: technical support with the system parameter identification.

6. Laschowski B, McNally W, Wong A, and McPhee J. (2020). Comparative analysis of environment recognition systems for control of lower-limb exoskeletons and prostheses. IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob). DOI: 10.1109/Bio Rob49111 .2020.9224364.

Contribution of W. McNally: technical support with the convolutional neural network evaluations.

7. Laschowski B, McNally W, Wong A, and McPhee J. (2020). ExoNet database: Wearable camera images of human locomotion environments. Frontiers in Robotics and AI, 7, 562061. DOI: 10. 3389/frobt.2020.562061.

Contribution of W. McNally: technical support with the code development for image labelling.

8. Laschowski B, Razavian RS, and McPhee J. (2021). Simulation of stand-to-sit biomechanics for robotic exoskeletons and prostheses with energy regeneration. IEEE Transactions on Medical Robotics and Bionics, 3(2), pp. 455-462. DOI: 10.1109/TMRB.2021.3058323.

Contribution of RS. Razavian: technical support with the system parameter identification.

9. Laschowski B, McNally W, Wong A, and McPhee J. (2021). Computer vision and deep learning for environment-adaptive control of robotic lower-limb exoskeletons. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). DOI: 10.1109/EMBC46164.2021.9630064.

Contribution of W. McNally: technical support with the code development for deep learning.

10. Nasr A, Laschowski B, and McPhee J. (2021). Myoelectric control of robotic leg prostheses and exoskeletons: A review. ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC-CIE). DOI: 10.1115/DETC2021 69203.

Contribution of B. Laschowski: assisted with the review of myoelectric control systems.

11. Laschowski B, McNally W, Wong A, and McPhee J. (2021). Environment classification for robotic leg prostheses and exoskeletons using deep convolutional neural networks. Frontiers in Neurorobotics. DOI: 10.3389/fnbot.2021.730965.

Contribution of W. McNally: technical support with the code development for deep learning.

ABSTRACT

Robotic leg prostheses and exoskeletons can provide powered locomotor assistance to older adults and/or persons with physical disabilities. However, limitations in automated control and energy-efficient actuation have impeded their transition from research laboratories to real-world environments.

With regards to control, the current automated locomotion mode recognition systems being developed rely on mechanical, inertial, and/or neuromuscular sensors, which inherently have limited prediction horizons (i.e., analogous to walking blindfolded). Inspired by the human vision-locomotor control system, here a multi-generation environment sensing and classification system powered by computer vision and deep learning was developed to predict the oncoming walking environments prior to physical interaction, therein allowing for more accurate and robust high-level control decisions. To support this initiative, the "ExoNet" database was developed – the largest and most diverse open-source dataset of wearable camera images of indoor and outdoor real-world walking environments, which were annotated using a novel hierarchical labelling architecture. Over a dozen state-of-the-art deep convolutional neural networks were trained and tested on ExoNet for large-scale image classification and automatic feature engineering. The benchmarked CNN architectures and their environment classification predictions were then quantitatively evaluated and compared using an operational metric called "NetScore", which balances the classification accuracy with the architectural and computational complexities (i.e., important for onboard real-time inference with mobile computing devices). Of the benchmarked CNN architectures, the EfficientNetB0 network achieved the highest test accuracy; VGG16 the fastest inference time; and MobileNetV2 the best NetScore. These comparative results can inform the optimal architecture design or selection depending on the desired performance of an environment classification system.

With regards to energetics, backdriveable actuators with energy regeneration can improve the energy efficiency and extend the battery-powered operating durations by converting some of the otherwise dissipated energy during negative mechanical work into electrical energy. However, the evaluation and control of these regenerative actuators has focused on steady-state level-ground walking. To encompass real-world community mobility more broadly, here an energy regeneration system, featuring mathematical and computational models of human and wearable robotic systems, was developed to simulate energy regeneration and storage during other locomotor activities of daily living, specifically stand-to-sit movements. Parameter identification and inverse dynamic simulations of subject-specific optimized biomechanical models were used to calculate the negative joint mechanical work and power while sitting down (i.e., the mechanical energy theoretically available for electrical energy regeneration). These joint mechanical energetics were then used to simulate a robotic exoskeleton being backdriven and regenerating energy. An empirical characterization of an exoskeleton was carried out using a joint dynamometer system and an electromechanical motor model to calculate the actuator efficiency and to simulate energy regeneration and storage with the exoskeleton parameters. The performance calculations showed that regenerating electrical energy during stand-to-sit movements provide small improvements in energy efficiency and battery-powered operating durations.

In summary, this research involved the development and evaluation of environment classification and energy regeneration systems to improve the automated control and energy-efficient actuation of next-generation robotic leg prostheses and exoskeletons for real-world locomotor assistance.

ACKNOWLEDGEMENTS

First and foremost, I want to thank Dr. John McPhee for his guidance and support. Over the years, John has helped me develop both personally and professionally. I also want to acknowledge my committee members - Drs. James Tung, Arash Arami, and Ning Jiang - for creating an environment at Waterloo Engineering that supports exoskeleton and prosthetics research. I also want to thank current and former members of the Motion Research Group for their encouragement and support, especially William McNally, Ali Nasr, Keaton Inkol, and Drs. Reza Razavian and Mahdokht Ezati. Special thanks to Dr. Alexander Wong for his mentorship in computer vision and deep learning, and to the Waterloo RoboHub and the Waterloo Artificial Intelligence Institute for providing a supportive research environment. I also want to thank the engineering staff at Technaid for their assistance with the robotic exoskeleton testing.

I also want to acknowledge those in the broader exoskeleton and prosthetics community for their personal correspondences and/or conversations at conferences, including Drs. Elliot Rouse; Levi Hargrove; Robert Gregg; Conor Walsh; Frank Sup; Tommaso Lenzi; Helen Huang; Amos Winter; Karl Zelik; Atakan Varol; and Michael Goldfarb. Special thanks to Dr. Jan Andrysek for his guidance during my tenure at the Holland Bloorview Kids Rehabilitation Hospital, where he first introduced me to prosthetics research. I want to thank the Natural Sciences and Engineering Research Council of Canada (NSERC), the Waterloo Engineering Excellence PhD Fellowship program, and the Canada Research Chairs (CRC) program for funding this research. I also want to recognize the TensorFlow Research Cloud program by Google and the NVIDIA GPU Grant program for providing the deep learning computing hardware.

DEDICATION

This thesis is dedicated to Katherine Laschowski, my wife and best friend.

TABLE OF CONTENTS

ABSTRACT ..... *vi*
LIST OF TABLES ..... *xii*
LIST OF FIGURES ..... *xv*
1. INTRODUCTION *1*
1.1 Motivation and Challenges *1*
1.1.1 Aging and Rehabilitation *1*
1.1.2 Passive Assistive Devices *1*
1.1.3 Robotic Prostheses and Exoskeletons *2*
1.1.4 Practical Limitations *4*
1.2 Thesis Outline *5*
2. LITERATURE REVIEW *7*
2.1 Automated Control *7*
2.1.1 Hierarchical Control Architecture *7*
2.1.2 Neuromuscular-Mechanical Sensing *7*
2.1.3 Environment Sensing *9*
2.2 Energy-Efficient Actuation *14*
2.2.1 Energetics of Human Locomotion *14*
2.2.2 Series Elastic Actuators *15*
2.2.3 Regenerative Actuators *17*
2.3 Chapter Summary *21*
3. ENVIRONMENT RECOGNITION *22*
3.1 Preliminary System Design *22*
3.1.1 Introduction *22*
3.1.2 Image Dataset of Walking Environments *22*
3.1.3 Convolutional Neural Network *23*
3.1.4 Classification Results *25*
3.2 Second-Generation System *28*
3.2.1 Introduction *28*
3.2.2 Large-Scale Image Dataset *28*
3.2.3 Deep Convolutional Neural Networks *34*
3.2.4 Classification Results *35*
3.3 Chapter Summary *45*

4. ENERGY REGENERATION 46
4.1 Biomechanics for Regeneration 46
4.1.1 Introduction 46
4.1.2 Motion Capture Experiments 48
4.1.3 Biomechanical Model 50
4.1.4 Simulation and Parameter Identification 51
4.1.5 Biomechanics Results 54
4.2 Regenerative Actuation 58
4.2.1 Introduction 58
4.2.2 Robotic Lower-Limb Exoskeleton 58
4.2.3 Dynamometer Testing 59
4.2.4 Actuator Efficiency 61
4.2.5 Regeneration Results 63
4.3 Chapter Summary 66
5. CONCLUSIONS AND FUTURE RESEARCH 67
5.1 Summary of Contributions 67
5.2 Recommendations for Future Research 69
5.2.1 Environment Recognition 69
5.2.1.1 Depth Sensing 69
5.2.1.2 Recurrent Neural Networks 70
5.2.1.3 Sensor Fusion 71
5.2.2. Energy Regeneration 72
5.2.2.1 Ultracapacitors 72
5.2.2.2 Actuator Design Optimization 73
5.2.2.3 Clinical Evaluations 74
REFERENCES 76

LIST OF TABLES

Table 2.1. Experimental datasets used for image classification of walking environments with applications to robotic leg prostheses and exoskeletons. Note that the datasets in Laschowski et al. [40] and [41] are from Sections 3.1 and 3.2 of this thesis, respectively. The dataset size is expressed as number of images. 13

Table 2.2. Environment recognition systems that used heuristics, statistical pattern recognition, or support vector machines for image classification of walking environments. Note that each classifier was developed and tested on different image datasets (see Table 2.1). The computation times are reported per image. 13

Table 3.1. Environment recognition systems that used convolutional neural networks for image classification of walking environments. Note that each neural network was trained and tested on different image datasets (Table 2.1). The computing hardware were mostly developed and manufactured by NVIDIA. The number of operations is expressed in multiply-accumulates. The inference times are reported per image. 27

Table 3.2. The benchmarked CNN architectures and their environment classification performances during inference on the ExoNet database. The test accuracies, parameters, and computing operations are expressed in percentages (0-100%), millions of parameters (M), and billions of multiply-accumulates (GMACs), respectively. Training and inference were both performed on a Google Cloud TPU. The EfficientNetB0 network achieved the highest test accuracy; VGG16 the fastest inference time; and MobileNetV2 the best NetScore and least number of parameters and computing operations. 37

Table 3.3. The multiclass confusion matrix for EfficientNetB0 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 37

Table 3.4. The multiclass confusion matrix for InceptionV3 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 38

Table 3.5. The multiclass confusion matrix for MobileNet showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 38

Table 3.6. The multiclass confusion matrix for MobileNetV2 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 38

Table 3.7. The multiclass confusion matrix for VGG16 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 39

Table 3.8. The multiclass confusion matrix for VGG19 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 39

Table 3.9. The multiclass confusion matrix for Xception showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 39

Table 3.10. The multiclass confusion matrix for ResNet50 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 40

Table 3.11. The multiclass confusion matrix for ResNet101 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 40

Table 3.12. The multiclass confusion matrix for ResNet152 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 40

Table 3.13. The multiclass confusion matrix for DenseNet121 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 41

Table 3.14. The multiclass confusion matrix for DenseNet169 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 41

Table 3.15. The multiclass confusion matrix for DenseNet201 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively. 41

Table 4.1. The average hip, knee, and ankle joint mechanical work (J/kg) per stride in healthy young adults (n=19) walking at natural cadence (110 ± 8 steps/min at 1.436 m/s) and normalized to total body mass [17]. “Total Work” is the combined mechanical energies from the hip, knee, and ankle joints and the “Net Joint Work” is the net mechanical work on each joint. Note that 42.6% of the total lower-limb joint mechanical work is negative (i.e., theoretically available for energy regeneration and storage). The results are for one leg. 63

Table 4.2. The average hip, knee, and ankle joint mechanical work (J/kg) per stand-to-sit movement in healthy young adults (~1 repetition per 2.25 seconds) and normalized to total body mass. The averages were taken across multiple subjects (n=9) and trials (20 trials/subject). “Total Work” is the combined mechanical energies from the hip, knee, and ankle joints and the “Net Joint Work” is the net mechanical work on each joint. Note that 91.2% of the total lower-limb joint mechanical work is negative (i.e., theoretically available for energy regeneration and storage). The results are for one leg. 63

Table 5.1. The average hip, knee, and ankle joint mechanical work (J/kg) per stride in older adults (n=18) walking at natural cadence (111 ± 8 steps/min at 1.28 m/s) and normalized to total body mass [57]. “Total Work” is the combined mechanical energies from the hip, knee, and ankle joints and the “Net Joint Work” is the net mechanical work on each joint. Note that 47.5% of the total lower-limb joint mechanical work is negative (i.e., theoretically available for energy regeneration and storage). The results are for one leg..... 75

LIST OF FIGURES

Figure 1.1 Examples of robotic leg prostheses under research and development. The photographs (left to right) were provided by Drs. Helen Huang (North Carolina State University and University of North Carolina at Chapel Hill, USA), Tommaso Lenzi (University of Utah, USA), Robert Riener (ETH Zurich, Switzerland), and Xiangrong Shen (University of Alabama, USA), respectively. 2

Figure 2.1. Hierarchical control architecture of robotic leg prostheses and exoskeletons, including high, mid, and low-level controllers. The high-level controller selects the desired locomotion mode using either (1) manual communication from the user (i.e., for commercially available devices) or (2) automated systems (i.e., for devices under research and development). 8

Figure 2.2. An automated locomotion mode recognition system for robotic leg prostheses and exoskeletons, also known as an intent recognition system or intelligent high-level controller. These systems can be supplemented with an environment recognition system to forward predict the oncoming walking environments prior to physical interaction, therein minimizing the high-level decision space. 10

Figure 2.3. Examples of walking with a robotic lower-lower exoskeleton with computer vision-based environment sensing superimposed. 12

Figure 2.4. The average hip, knee, and ankle joint mechanical power (W/kg) per stride in healthy young adults (n=19) walking at natural cadence (110 ± 8 steps/min at 1.44 m/s) and normalized to total body mass (top left). The positive and negative values represent joint mechanical power generation and absorption, respectively. Data were taken from Winter [54], the trajectories of which begin and end with heel-strike (top right). These locomotor energetics have implications on energy-efficient actuation of robotic leg prostheses and exoskeletons during motoring and braking operations (bottom left); the nomenclature are described in the text. 16

Figure 3.1. Schematic of the preliminary environment recognition system powered by computer vision and deep learning. Images of real-world walking environments were collected using a lightweight wearable camera (1) and classified using a 10-layer convolutional neural network (2). The classification predictions during inference for each environment class are displayed using a multiclass confusion matrix (3). 24

Figure 3.2. Examples of failure cases whereby the convolutional neural network incorrectly predicted the walking environment. Images in the top, middle, and bottom rows were misclassified as incline stairs, level-ground, and decline stairs, respectively. 26

Figure 3.3. Examples of convolutional neural networks used for image classification of walking environments. The top and bottom schematics were adapted from Drs. Dan Simon (Cleveland State University, USA) and Chenglong Fu (Southern University of Science and Technology, China), respectively. 27

Figure 3.4. Development of the "ExoNet" database, including (1) a photograph of the wearable camera system used for large-scale data collection; (2) examples of the high-resolution RGB

images of walking environments; and (3) a schematic of the novel hierarchical labelling architecture with 12 classes. 29

Figure 3.5. Examples of the wearable camera images of indoor and outdoor real-world walking environments in the ExoNet database. Images were collected at various times throughout the day and across different seasons (i.e., summer, fall, and winter). 31

Figure 3.6. Schematic of the class distributions in the ExoNet database, the images of which were annotated using a novel hierarchical labelling architecture. A description of the class labels is provided in the text. 32

Figure 3.7. Examples of "steady" and "transition" states in the ExoNet hierarchical labelling architecture. The images in the top and bottom rows were labelled as steady states and those in middle row were labelled as transition states. For each column, the left images show the author walking with a robotic exoskeleton and the right images show the concurrent field-of-view of the wearable camera system (i.e., what the exoskeleton would see). 33

Figure 3.8. The loss and image classification accuracies during training and validation on the ExoNet database using state-of-the-art deep convolutional neural networks, including: EfficientNetB0, InceptionV3, MobileNet, MobileNetV2, VGG16, VGG19, Xception, ResNet50, ResNet101, ResNet152, DenseNet121, DenseNet169, and DenseNet201. 36

Figure 4.1. Examples of the stand-to-sit movements used for energy regeneration, the biomechanics of which were experimentally measured using motion capture cameras and force plates. Nine healthy young adults (n=9) were tested in total. 49

Figure 4.2. Schematic of the experimental and computational methods used to evaluate the stand-to-sit biomechanics for energy regeneration, including (1) the motion capture experiments; (2) biomechanical model; (3) dynamic parameter identification; (4) joint power analyses; and (5) a regenerative actuator model used to convert human mechanical power $\tau j\theta j$ into electrical power iv while sitting down. 52

Figure 4.3. The percent difference in HAT, thigh, and shank body segment mass between the system parameter identification (ID) and anthropometric proportions from cadaver research [59] scaled to the measured total body mass of each subject. The positive and negative values represent an overestimation and underestimation in body segment mass, respectively, from the parameter ID compared to the anthropometric proportions. The upper and lower bounds in the optimization are also displayed, the values of which are described in the text. 53

Figure 4.4. The average hip, knee, and ankle joint angles (°) and torques (Nm/kg) during stand-to-sit movements from inverse kinematics and dynamics, respectively. The joint torques are normalized to total body mass and are reported per leg. Subjects were healthy young adults. The uncertainties are ± one standard deviation across multiple subjects (n=9) and trials (20 trials/subject). 55

Figure 4.5. The average hip, knee, and ankle joint mechanical power (W/kg) per leg during level-ground walking at natural cadence (n=19) and stand-to-sit movements (n=9), normalized to total body mass. Subjects were all healthy young adults. The positive and negative values represent

joint power generation and absorption, respectively. The walking data were taken from Winter [54], the trajectories of which begin and end with heel-strike. The uncertainties are ± one standard deviation across multiple subjects and trials..... 57

Figure 4.6. Schematic of the exoskeleton benchtop testing using a joint dynamometer system to calculate the actuator efficiency. The exoskeleton knee was kinematically driven through sinusoidal flexion and extension movements while the dynamometer measured the joint torque and angular speed (i.e., the mechanical power output). The onboard exoskeleton sensors, and an electromechanical motor model, were used to estimate the motor voltage and current (i.e., the electrical power input). The actuator efficiency is the ratio of the average instantaneous power outputs to inputs over the steady-state time intervals. The nomenclature are described in the text. Details regarding the exoskeleton control architecture were taken from [145]..... 60

Figure 5.1. The average hip, knee, and ankle joint mechanical power (W/kg) per leg while walking at natural cadence in young adults (n=19) and older adults (n=18), normalized to total body mass. The positive and negative values represent joint power generation and absorption, respectively. Data were taken from Winter [57], the trajectories of which begin and end with heel-strike. The uncertainties are ± one standard deviation across multiple subjects and trials. 75

1. INTRODUCTION

1.1 Motivation and Challenges

1.1.1 Aging and Rehabilitation

There are currently hundreds of millions of individuals worldwide with mobility impairments resulting from aging and/or physical disabilities, the numbers of which are expected to significantly increase over the coming decades. For instance, the number of persons with limb amputations in the United States alone is expected to more than double from ~1.6 million in 2005 to ~3.6 million in 2050 [1]. This increased prevalence is primarily driven by the aging population and higher rates of dysvascular diseases, especially diabetes, among older adults. The World Health Organization recently projected that the number of individuals above 65 years age will increase from ~524 million in 2010 to ~1.5 billion in 2050 [2]. The global prevalence of musculoskeletal disorders like osteoarthritis is ~151 million and neurological conditions like Parkinson's disease, cerebral palsy, and spinal cord injury affect approximately 5.2 million, 16 million, and 3.5 million persons, respectively [2]. Furthermore, the number of individuals needing physical rehabilitation due to immobility and/or disease has recently been exacerbated by the coronavirus disease 2019 (COVID-19) [3]. Mobility impairments, especially those resulting in wheelchair dependency, are often associated with reduced independence and secondary health complications, including osteoporosis, coronary artery disease, obesity, hypertension, and sarcopenia [2]. Consequently, the demand for mobility assistive technologies is dramatically increasing.

1.1.2 Passive Assistive Devices

Fitted with conventional passive assistive devices, persons with mobility impairments tend to fall more frequently and walk more energetically inefficient and with greater biomechanical asymmetries compared to able-bodied individuals, which have implications on joint and bone degeneration resulting from disproportionate loading on the unaffected limb over time, assuming unilateral impairments [4]–[6]. For example, persons with transtibial and transfemoral amputations expend up to 30% and 60% more metabolic energy, respectively, compared to non-amputees walking at the same speed, and exhibit 10-65% slower self-selected walking speeds [5]. One reason for these gait abnormalities is that many locomotor activities require net positive mechanical work about the lower-limb joints via power generation from the human muscles. However, limb amputation or paralysis removes/diminishes these sources of biological actuation, therein compromising the user's ability to perform many locomotor activities effectively and efficiently. Passive assistive devices (e.g., the prosthetic leg used by Terry Fox, the Canadian humanitarian and cancer research activist) cannot generate positive mechanical power and lack the ability to actively adapt to the human and/or changes in the walking environment. These

devices can only passively control the leg motion using mechanical elements in induce friction or hydraulic or pneumatic dampers.

1.1.3 Robotic Prostheses and Exoskeletons

Fortunately, newly-developed robotic leg prostheses and exoskeletons can replace the propulsive function of the amputated or impaired biological muscles and allow individuals with mobility impairments to perform daily locomotor activities that require positive power generation using motorized hip, knee, and/or ankle joints [4]–[8]. Generally speaking, these wearable robotic devices feature biomimetic design principles, whereby the actuators and mechanical structure mimic the human musculoskeletal system; the sensors and microcontroller mimic the peripheral and central nervous systems, respectively; and the batteries mimic the metabolic power sources. Figure 1.1 shows several examples of robotic leg prostheses under research and development, which operate in series with the biological limb.

With regards to robotic exoskeletons, which operate in parallel with the musculoskeletal system, there are three broad categories of devices: 1) human performance augmentation exoskeletons designed to increase the physical strength and endurance of able-bodied individuals; 2) assistive exoskeletons that help older adults and/or persons with physical disabilities perform daily locomotor activities; and 3) rehabilitation exoskeletons, which are mainly treadmill-based systems (e.g., Lokomat), designed to promote motor learning and improve the user's physical

Figure 1.1 Examples of robotic leg prostheses under research and development. The photographs (left to right) were provided by Drs. Helen Huang (North Carolina State University and University of North Carolina at Chapel Hill, USA), Tommaso Lenzi (University of Utah, USA), Robert Riener (ETH Zurich, Switzerland), and Xiangrong Shen (University of Alabama, USA), respectively.

capabilities when not using the device. The proliferation of exoskeleton technologies largely began around the early 2000s when the Defense Advanced Research Projects Agency (DARPA) funded the development of human performance augmentation exoskeletons; one of the most notable prototypes was the hydraulically-powered Berkeley Lower-Extremity Exoskeleton (BLEEX) device developed by Kazerooni's lab at the University of California, Berkeley (USA) [9].

The ReWalk device is one of the oldest and most established commercial exoskeletons for assisting persons with physical disabilities, specifically spinal cord injury [8]. The exoskeleton bilaterally actuates the hip and knee joints for walking and sit-to-stand motions. The user controls the high-level function with a remote control and carries the batteries and controller in a backpack. In comparison, the Rex Bionics exoskeleton completely encompasses both lower-limbs and the user stands on foot platforms while the machine physically walks for them with minimal human-robot interaction. Several researchers have used the Rex Bionics exoskeleton to integrate brain-machine interfaces for neural control [8]. The Indego powered exoskeleton, designed for persons with spinal cord injury, bilaterally actuates the hip and knee joints and includes a waist-mounted battery pack. Because the Indego device originated from Goldfarb's lab at Vanderbilt University (USA), more information is publicly available about its mechatronic design and control [10]. The joint controller functions in either position control, which sets high gains to enforce predefined kinematic trajectories, or impedance control, which virtually emulates a spring-damper system. The high-level controller includes individual finite state machines, and the user manually switches between different states depending on the locomotion mode. In addition to robotic exoskeletons, the Össur Power Knee, the only commercially-available powered knee prosthesis, provides locomotor assistance using an electrically-motorized knee and manual high-level control based on exaggerated movements performed by the user [5].

Although early robotic leg prostheses and exoskeletons relied on hydraulic and pneumatic actuators tethered to off-board fluid pumps and air compressors, respectively, the field has shifted towards using electromagnetic actuators for power generation, specifically brushed and brushless direct current (DC) motors [4]–[8]. Electric motors tend to be most efficient at low torques and high speeds, with torque and power densities around 15 Nm/kg and 200 W/kg, respectively, with an efficiency ~90% [11]. In comparison, human muscles have torque and power densities around 20 Nm/kg and 50 W/kg, respectively, with an efficiency ~30% during concentric contractions [11]. Most electric motors are coupled with a high-ratio transmission (e.g., ball-screw mechanism or harmonic gearing) to increase the motor torque output to that needed for legged locomotion; this design causes the robotic actuator to have high output impedance (i.e., mechanically stiff), which allows for precision position control [12]. For instance, the ReWalk and Ekso Bionics powered exoskeletons use stiff actuators to rigidly track predefined kinematic trajectories, which can benefit users with limited ability to physically interact with and control the robotic device (e.g., persons with complete paralysis) [8]. However, actuators with high output impendence cannot

exploit the passive dynamics of legged locomotion and/or other energy storage and return mechanisms, therein often resulting in higher energy consumption and peak power requirements of the electric motor [6], [12].

1.1.4 Practical Limitations

Despite their ability to replace the propulsive function of the amputated or impaired biological muscles and generate positive mechanical power using motorized joints, robotic leg protheses and exoskeletons still face many practical challenges limiting their widespread adoption in real-world environments. With regards to control, high-level transitions between different locomotor activities remain a significant challenge. Inaccurate and/or delayed control decisions could result in loss-of-balance and injury, which can be especially problematic when involving stairs. For commercial devices, switching between different mode-specific controllers is supervised by a high-level controller based on direct communication from the user. For example, the Össur Power Knee prosthesis and the Indego and ReWalk powered exoskeletons require the user to perform exaggerated movements or use hand controls to manually switch between locomotion modes [4], [8]. Although highly accurate in communicating the user's locomotor intent to the robotic device, manual high-level control and decision-making can be time-consuming, inconvenient, and cognitively demanding [13]. Transitions between different locomotor activities should ideally be seamless and intuitive without requiring volitional action from the user. Researchers have been working on developing automated locomotion mode recognition systems to shift the high-level control burden from the user to an intelligent controller, as reviewed in Section 2.1.

With regards to energetics, limitations in battery and actuator technologies have impeded the untethered operation of robotic leg prostheses and exoskeletons in real-world environments. Traditional rigid actuators used in robotics tend to be heavy and inefficient, which can increase the power consumption and thus decrease the operating durations and/or require larger onboard batteries [7], [8]. For example, robotic knee prostheses under research and development weigh 2-5 kg and provide only 3 ± 2 hours of operating time [14]. Similarly, most robotic exoskeletons provide 1-5 hours of maximum battery-powered operation [8]. Onboard portable power has often been considered one of the leading challenges to developing robotic exoskeletons for real-world environments [7], [8]. Increased device mass and inertia, due to heavier onboard motors and batteries, would require more effort by the human musculoskeletal system during swing phase, thus reducing the locomotor efficiency via higher metabolic power consumption [15]. For socket-suspended prostheses, increased device weight could also cause pain and discomfort due to greater tensile forces on the human-prosthesis interface [16], [17]. Consequently, many review papers [4]–[8] have concluded that further advances in energy-dense batteries and lightweight and efficient actuators are needed to support the untethered operation of robotic leg prostheses and exoskeletons, as reviewed in Section 2.2.

1.2 Thesis Outline

Motivated by these practical limitations, the overall objective of this thesis research is to support the energy-efficient actuation and automated control of next-generation robotic leg prostheses and exoskeletons for real-world locomotor assistance. Chapter 2 reviews the state-of-the-art in 1) locomotion mode recognition systems used for automated high-level control and decision making; and 2) energy-efficient actuators, including series elastic actuators and/or backdriveable actuators with energy regeneration, which exploit the energetics of human legged locomotion.

Chapter 3 presents the development and evaluation of a multi-generation environment recognition system powered by deep learning and computer vision to predict the oncoming walking environment prior to physical interactions, therein allowing for more accurate and robust automated high-level control decisions. This includes 1) the development of the largest and most diverse open-source dataset of wearable camera images of real-world walking environments, known as the "ExoNet" database; 2) the training and testing of state-of-the-art deep convolutional neural networks (CNNs) on the ExoNet dataset for image classification and automatic feature engineering; and 3) comparison of the benchmarked CNN architectures and their environment classification predictions using an operational metric conducive to onboard real-time inference. This environment sensing and classification system provides a large-scale benchmark for future research in environment-adaptive locomotor control.

Chapter 4 presents the development and evaluation of an energy regeneration system, featuring mathematical and computational models of human and wearable robotic systems, to simulate energy regeneration and storage during locomotor activities of daily living, specifically stand-to-sit movements. Parameter identification and inverse dynamic simulations of subject-specific optimized biomechanical models were used to calculate the negative joint mechanical work and power while sitting down (i.e., the mechanical energy theoretically available for electrical energy regeneration). These lower-limb joint mechanical energetics were then used to simulate backdriving a robotic exoskeleton and regenerating energy. An empirical characterization of an exoskeleton was carried out using a joint dynamometer system and an electromechanical motor model to calculate the actuator efficiency and to simulate energy regeneration and storage with the exoskeleton parameters. The performance calculations show that regenerating electrical energy during stand-to-sit movements provide small improvements in energy efficiency and battery-powered operating durations.

Building on this thesis research, Chapter 5 recommends future directions for research and innovation. With regards to environment recognition, these include 1) the application of depth cameras for 3D environment sensing; 2) the classification of sequential walking images using recurrent neural networks; and 3) multi-sensor data fusion for automated locomotion mode recognition. With regards to energy regeneration, these include 1) the application of advanced

ultracapacitors for onboard energy storage; 2) the optimization of actuator design parameters for energy efficiency and performance; and 3) the study of older adults and/or persons with physical disabilities to evaluate the potential for electrical energy regeneration. In addition to robotic leg prostheses and exoskeletons, the principles of environment classification and energy regeneration could extend to humanoids, autonomous legged robots, powered wheelchairs, and other mobility assistive technologies.

2. LITERATURE REVIEW

To motivate this thesis research, the following chapter reviews the state-of-the-art in locomotion mode recognition systems used for automated high-level control and decision making (see Section 2.1); and energy-efficient actuators, including series elastic actuators and/or backdriveable actuators with energy regeneration, which exploit the energetics of human legged locomotion (see Section 2.2).

2.1 Automated Control

2.1.1 Hierarchical Control Architecture

Most robotic leg prostheses and exoskeletons use a hierarchical control architecture, including high, mid, and low-level controllers [4] (Figure 2.1). The high-level controller is responsible for state estimation and predicting the user's locomotor intent. The mid-level controller converts the locomotor activity into mode-specific reference trajectories using dynamic equations of the biomechatronic system; this level of control typically includes individual finite-state machines with discrete mechanical impedance parameters like stiffness and damping coefficients, which are heuristically tuned for different locomotor activities to generate the desired actuator joint torques (i.e., the desired device state). Tunning these parameters can be time-consuming as the number of parameters increases with the number of states per locomotor activity, the number of activities, and the number of actuated joints. The low-level controller uses standard controls algorithms like proportional-integral-derivative (PID) control to calculate the error between the measured and desired device states and command the robotic actuators to minimize the error using reference tracking and closed-loop feedback control. Unlike commercial devices that require direct communication from the user for manual high-level control and decision-making, researchers have been working on developing automated locomotion mode recognition systems that use pattern recognition algorithms and data from wearable sensors like inertial measurement units (IMUs) and surface electromyography (EMG).

2.1.2 Neuromuscular-Mechanical Sensing

Mechanical sensors embedded in robotic leg prostheses and exoskeletons (e.g., potentiometers, pressure sensors, magnetic encoders, and strain gauge load cells) can be used for state estimation by measuring the joint angles and angular velocities, and interaction forces and/or torques between the human and device, and between the device and environment [4]. These sensors resemble how biological proprioceptors and mechanoreceptors provide kinematic and force feedback to the central nervous system, respectively. For example, Varol and colleagues [18] used onboard mechanical sensors and a Gaussian mixture model classification algorithm for automated high-level control of a robotic leg prosthesis. In addition to mechanical sensors, inertial

measurement units can measure 3D angular velocities, accelerations, and direction of lower-limb segments using an onboard gyroscope, accelerometer, and magnetometer, respectively. Although mechanical and inertial sensors allow for fully integrated control systems, such that no additional instrumentation or wiring apart from the device need be worn, these sensors can only respond to the user's movements.

In contrast, the electrical potentials of biological muscles, as experimentally recorded using surface electromyography, precede movement initiation and thus could predict locomotion mode transitions with small prediction horizons. Some common leg muscles used for prosthetic control include the semitendinosus, adductor magnus, sartorius, vastus medialis and lateralis, biceps and rectus femoris, and gracilis [14]. For example, using surface EMG and a linear discriminant analysis

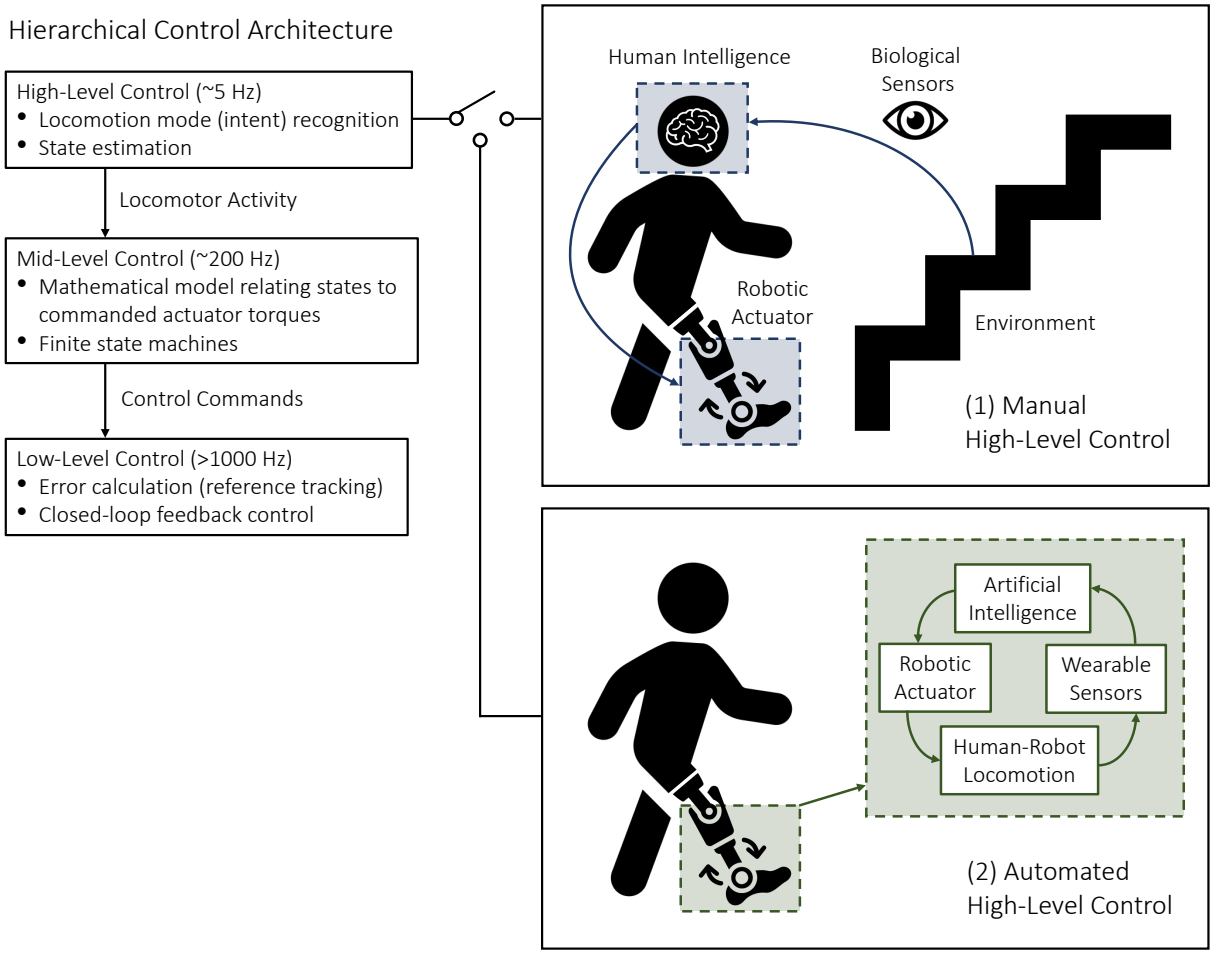

Figure 2.1. Hierarchical control architecture of robotic leg prostheses and exoskeletons, including high, mid, and low-level controllers. The high-level controller selects the desired locomotion mode using either (1) manual communication from the user (i.e., for commercially available devices) or (2) automated systems (i.e., for devices under research and development).

(LDA) pattern recognition algorithm, Huang et al. [19] differentiated between seven locomotion modes with ~90% classification accuracy. In addition to automated locomotion mode recognition, EMG signals could be used for proportional myoelectric control [16], [20], [21]. Although several researchers have explored using brain-machine interfaces for direct neural control of robotic prostheses and exoskeletons (i.e., using either implanted or electroencephalography-based systems) [22], these sensor technologies are still highly experimental and infrequently used for locomotor applications.

Fusing information from mechanical and/or inertial sensors with surface EMG, known as neuromuscular-mechanical data fusion, can improve the locomotion mode recognition accuracies and decision times compared to implementing either system individually [23]–[29]. Huang et al. [27] first demonstrated such improvements in classification predictions across six locomotion modes using a support vector machine (SVM) classifier, surface EMG data, and measured ground reaction forces and torques on a robotic leg prosthesis. However, neuromuscular-mechanical data are user-dependent, therein often requiring time-consuming experiments to amass individual datasets, and surface EMG require calibration and are susceptible to fatigue, motion artifacts, changes in electrode-skin conductivity, and crosstalk between adjacent muscles [4], [8]. EMG signals may also be inaccessible in persons with high amputations where insufficient neural content is available from the residual limb. Despite advances in automated intent recognition using mechanical, inertial, and/or neuromuscular signals, further improvements in performance are desired for safe and robust locomotor control. Neuromuscular-mechanical sensors also have limited prediction horizons such that the data can only characterize the current state of the user and/or device, which is somewhat analogous to walking blindfolded.

2.1.3 Environment Sensing

Taking inspiration from the human vision-locomotor control system, supplementing neuromuscular-mechanical data with information about the oncoming walking environment could improve the automated high-level control performance (Figure 2.2). Environment sensing would precede modulation of the user's muscle activations and/or walking biomechanics, therein allowing for more accurate and robust transitions between locomotion modes by minimizing the high-level decision space. During human legged locomotion, supraspinal levels of the central nervous system acquire state information from biological sensors like the eyes through ascending pathways, which are then used to actuate and control the musculoskeletal system through feedforward efferent commands [30]. However, these control loops are compromised in persons using assistive devices due to limitations in the human-machine interaction. Environment sensing and classification could artificially restore these control loops for automated high-level control and decision making. Environment information could also be used to adapt the mid-level reference

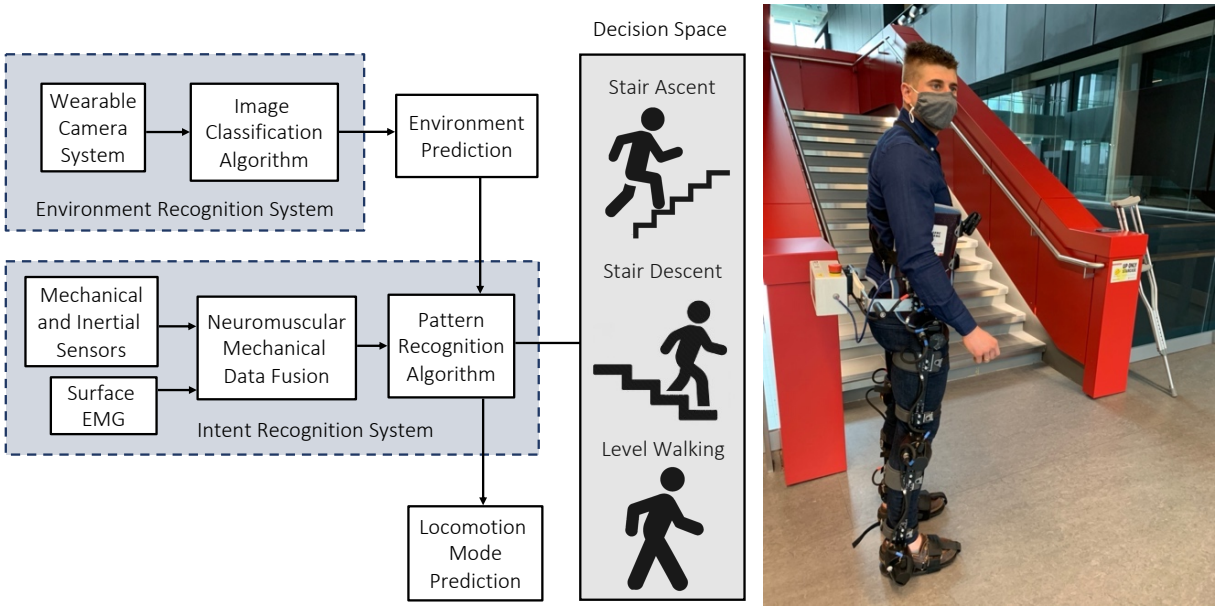

Figure 2.2. An automated locomotion mode recognition system for robotic leg prostheses and exoskeletons, also known as an intent recognition system or intelligent high-level controller. These systems can be supplemented with an environment recognition system to forward predict the oncoming walking environments prior to physical interaction, therein minimizing the high-level decision space.

trajectories (e.g., increasing actuator joint torques for toe clearance corresponding to an obstacle height) [31]; optimal path planning (e.g., identifying opportunities for energy recycling) [32], [33]; and varying foot placement based on the walking surface.

Some of the earliest research fusing neuromuscular-mechanical data with environment information for prosthetic leg control came from Huang's research lab at the University of North Carolina at Chapel Hill and North Carolina State University [23], [24], [26], [28], [34]. Different walking environments were statistically modelled as prior probabilities using the principle of maximum entropy and incorporated into the discriminant function of an LDA classification algorithm. Rather than using equal prior probabilities, the researchers simulated different walking environments by adjusting the prior probabilities of each class, which allowed their locomotion mode recognition system to dynamically adapt to different environments. For instance, when approaching an incline staircase, the prior probability of performing stair ascent would progressively increase while the prior probabilities of other locomotion modes would likewise decrease. Using these adaptive prior probabilities based on the terrain information significantly outperformed (95.5% classification accuracy) their locomotion mode recognition system based on neuromuscular-mechanical data alone with equal prior probabilities (90.6% accuracy) [24]. These seminal studies showed 1) how environment information could be incorporated into an intelligent high-level controller; 2) that including such information could improve the automated locomotion

mode recognition accuracies and decision times; and 3) that the controller could be relatively robust to noisy and imperfect environment predictions such that the neuromuscular-mechanical data dominated the high-level decision making.

Several researchers have explored using radar detectors [35] and laser rangefinders [24], [26], [34] for active environment sensing. Unlike camera-based systems, these sensors circumvent the need for computationally expensive image processing and classification. Radar can uniquely measure distances through non-conducting materials like clothing and are invariant to outdoor lighting conditions and surface textures. Using a leg-mounted radar detector, the Biomechatronics group at MIT [35] were able to measure stair distances and heights within ~1.5 cm and ~0.4 cm average accuracies, respectively, up to 6.3 m maximum distances. However, radar reflection signatures struggle with source separation of multiple objects and have relatively low resolution. Huang and colleagues [24], [26], [34] developed a waist-mounted system using an IMU and laser rangefinder to reconstruct the geometry of the oncoming walking environments between 300-10,000 mm ranges. Environmental features like the terrain height, distance, and slope were used for classification using heuristic rule-based thresholds. The system achieved 98.1% classification accuracy [34]. Although simple and effective, their system required subject-specific calibration (e.g., the device mounting height) and provided only a single distance measurement.

Compared to radar and laser rangefinders, cameras can provide more detailed information about the sampled field-of-view and detect physical obstacles and terrain changes in peripheral locations (see Figure 2.3). Most environment recognition systems for robotic leg prostheses and exoskeletons have used either RGB cameras [36]–[43] or 3D depth cameras [29], [31], [44]–[50] mounted on the chest [40], [41], [46], waist [29], [38], [47], [48], or leg [31], [36], [37], [43]–[45], [49]–[51] (Table 2.1). For example, Tung's research lab [52], [53] at the University of Waterloo (Canada) used a chest-mounted camera to collect images of slope changes (e.g., stairs, curbs, and ramps) and surfaces (e.g., gravel, grass, and concrete). However, their images were only grey-scale and the dataset was relatively small in size, which have implications for data-driven algorithms in order to prevent overfitting and promote generalization. Few studies have used head-mounted cameras for biomimicry [42], [43], [54]. Zhong et al. [43], [51] compared the effects of different wearable camera positions on environment sensing and classification. Compared to glasses, their leg-mounted camera more accurately detected closer walking environments but struggled with incline stairs, often capturing only 1-2 steps. Although the glasses could detect further-away environments, the head-mounted camera also captured irrelevant features like the sky, which affected the classification accuracy. The glasses also struggled with detecting decline stairs and had larger standard deviations in predictions due to head movement.

For image classification, researchers have traditionally used statistical pattern recognition and machine learning algorithms like support vector machines, which require hand-engineering (Table 2.2) [29], [36], [37], [39], [44]–[47], [52]. For example, Hargrove's research group at the

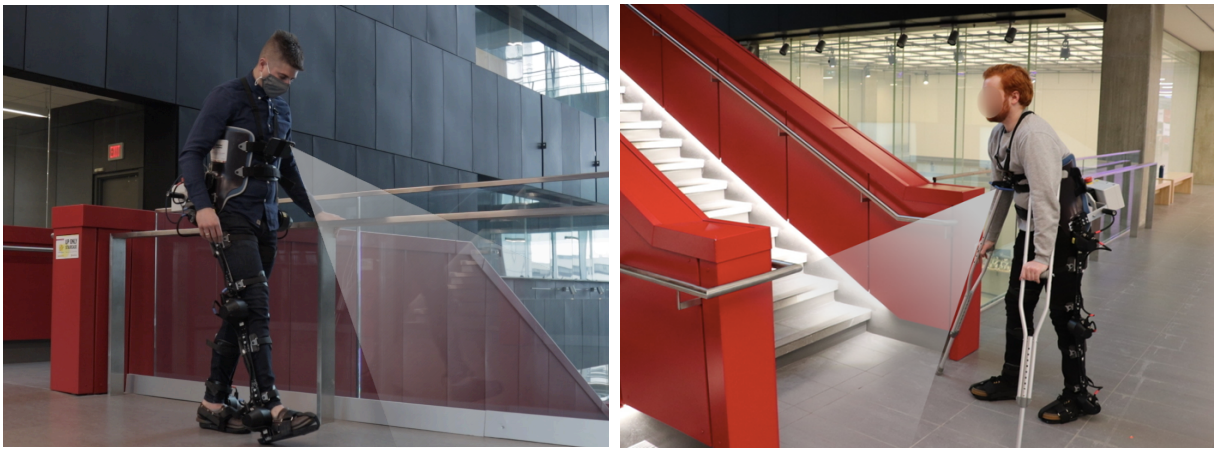

Figure 2.3. Examples of walking with a robotic lower-lower exoskeleton with computer vision-based environment sensing superimposed.

Rehabilitation Institute of Chicago and Northwestern University [39], [46] used standard image processing and rule-based thresholds to detect convex and concave edges and vertical and horizontal planes for stair recognition. Although their algorithm achieved 98.8% classification accuracy, the computations were time-consuming (~8 frames/sec) and the system was evaluated using only five images. Huang and colleagues [37] achieved 86% image classification accuracy across six environment classes using SURF features and a bag-of-words classifier.

Varol's research lab at Nazarbayev University (Kazakhstan) used support vector machines for classifying depth images of walking environments [44], [45]. These supervised machine learning algorithms work by mapping extracted features into a high-dimensional feature space and separate samples into different classes by constructing optimal hyperplanes with maximum margins. Different dimensionality reductions and SVM models were tested. Their system achieved 94.1% classification accuracy across five locomotion modes using a cubic kernel SVM and no dimension reduction [44]. Although support vector machines are effective in high-dimensional space, these algorithms require manual selection of kernel functions and statistical features, which can be time-consuming and suboptimal. In contrast, deep learning replaces hand-designed features with multilayer networks that can automatically and efficiently learn the optimal image features from training data. Convolutional neural networks also tend to outperform support vector machines for image classification [55]. These benefits of deep learning over traditional image classification algorithms motivated the thesis research in Chapter 3. This review of environment sensing and classification systems for robotic leg prostheses and exoskeletons was published in the 2020 IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob) [56].

Table 2.1. Experimental datasets used for image classification of walking environments with applications to robotic leg prostheses and exoskeletons. Note that the datasets in Laschowski et al. [40] and [41] are from Sections 3.1 and 3.2 of this thesis, respectively. The dataset size is expressed as number of images.

Reference	Camera	Body Position	Dataset Size	Image Resolution	Classes
Da Silva et al. [36]	RGB	Leg	3,992	512×512	6
Diaz et al. [37]	RGB	Leg	3,992	1080x1920	6
Khademi and Simon [38]	RGB	Waist	7,284	224x224	3
Krausz and Hargrove [39]	RGB	Head	5	928x620	2
Krausz et al. [46]	Depth	Chest	170	80x60	2
Krausz et al. [47]	Depth	Waist	4,000	171x224	5
Laschowski et al. [40]	RGB	Chest	34,254	224x224	3
Laschowski et al. [41]	RGB	Chest	922,790	1280×720	12
Massalin et al. [44]	Depth	Leg	402,403	320x240	5
Nouredanesh et al. [52]	RGB	Chest	12,382	32x32	17
Nouredanesh et al. [53]	RGB	Chest	3,669	32x32	12
Novo-Torres et al. [42]	RGB	Head	40,743	128x128	2
Rai et al. [54]	RGB	Head	N/A	640x480	3
Varol and Massalin [45]	Depth	Leg	22,932	320x240	5
Zhang et al. [49], [50]	Depth	Leg	7,500	224x171	5
Zhang et al. [48]	Depth	Waist	4,016	2048-point cloud	3
Zhang et al. [31]	Depth	Leg	7,500	100x100	5
Zhong et al. [43], [51]	RGB	Head and leg	327,000	1240x1080	6

Table 2.2. Environment recognition systems that used heuristics, statistical pattern recognition, or support vector machines for image classification of walking environments. Note that each classifier was developed and tested on different image datasets (see Table 2.1). The computation times are reported per image.

Reference	Feature Extractor and Classifier	Computing Devices	Test Accuracy (%)	Computation Time (ms)
Da Silva et al. [36]	Local binary pattern and random forest	NVIDIA Jetson TX2	90.0	200
Diaz et al. [37]	SURF features and bag-of-words model	Intel Core i7-2600 CPU (3.40GHz)	86.0	N/A
Krausz and Hargrove [39]	Hough transform with Gabor filter or canny edge detector	Intel Core i5	N/A	8000
Krausz et al. [46]	Heuristic thresholding and edge detector	Intel Core i5	98.8	200
Krausz et al. [47]	Regions-of-interest and linear discriminate analysis	Intel Core i7-8750H (2.2GHz)	N/A	N/A
Massalin et al. [44]	Cubic kernel support vector machine	Intel Core i7-2640M (2.8GHz)	94.1	14.9
Nouredanesh et al. [52]	Gabor barcodes and Hamming distances	Intel Core i7 (3.60GHz)	88.5	150
Varol and Massalin [45]	Support vector machine	Intel Core i7-2640M (2.8GHz)	99.0	14.9

2.2 Energy-Efficient Actuation

With regards to energetics, traditional robotic leg prostheses and exoskeletons have not exploited the passive dynamics of legged locomotion and/or other energy storage and return mechanisms, therein leading to heavy and inefficient actuators that require significant power consumption and thus provide limited battery-powered operating durations given the finite energy density of batteries. To address these limitations, researchers have been working on designing lightweight and efficient actuators that more effectively utilize the energetics of legged locomotion, including series elastic actuators and/or backdriveable actuators with energy regeneration, as subsequently reviewed. Given that the biomechanics of human walking is fundamental to the design of optimal energy-efficient actuators for wearable robotic devices, the energetics of legged locomotion are first discussed.

2.2.1 Energetics of Human Locomotion

Joint mechanical power is defined as the product of the net joint torque and angular velocity, and joint mechanical work is the cumulative time-integral of the joint mechanical power. During energy generation, the net joint torque and angular velocity have the same sign direction and positive mechanical work is done (e.g., a concentric contraction wherein the biological muscles shorten under tension). During energy absorption, the net joint torque and angular velocity have opposite polarities and negative mechanical work is done (e.g., an eccentric contraction wherein the biological muscles lengthen under tension); this assumes that the joint torque generators are independent of adjacent joints such that biarticulating muscles spanning multiple joints are ignored. The net rate of energy generation or absorption by all muscles crossing the joint is the joint mechanical power. During walking, some mechanical energy can be recycled by conservative forces (e.g., the elastic storage and return of muscle-tendon units or the pendular dynamics of swinging limbs) and transferred between adjacent segments [57]. Most models of legged locomotion ignore the elastic potential energy of deformable segments since the amount of deformation is relatively small and difficult to measure [57]–[59].

The energetics of human walking can be modelled by the mechanical work and power done on the total body system [60]–[62]. During single support, the stance leg resembles an inverted pendulum, whereby no net mechanical work is needed to move the center of mass (COM) and energy is conserved. During step-to-step transitions, however, external mechanical work mainly by ground reaction forces is needed to redirect the body's COM velocity from one pendulum arc to another, which is a major determinant of the metabolic cost of human locomotion [62]. To maintain steady-state level-ground walking, the leading leg performs negative mechanical work to redirect the COM velocity at foot-ground contact, while the trailing leg simultaneously performs positive mechanical work during push-off to restore the lost energy [60], [61]. For example, when walking at 1.25 m/s, 15.4 $\pm$ 2.6 J of positive external mechanical work is performed by the trailing

leg and 12.4 ± 3.1 J of negative external mechanical work is performed by the leading leg [62]. In theory, the net mechanical work during level-ground walking at constant speed should be nearly zero since there is no net change in the gravitational potential energy or translational kinetic energy of the total body system. Compared to external mechanical work done on the center of mass, joint mechanical work can more accurately model the human musculotendon work [63].

Winter at the University of Waterloo (Canada) analyzed the lower-limb joint mechanical work and power during level-ground walking using inverse dynamics [57]–[59] (see Figure 2.4). Generally speaking, the knee joint behaves like a damper mechanism, performing net negative mechanical work with four main power phases: 1) negative mechanical power at weight acceptance wherein the knee flexes under the control of an extensor moment; 2) positive mechanical power generation by the knee extensors during mid-stance such that the product of the extensor moment and angular velocity is positive; 3) negative mechanical power absorption by the extensors as the knee flexes during early swing; and 4) negative mechanical power absorption by the knee flexors during late swing to decelerate leg extension prior to heel-strike. In comparison, the ankle generally behaves like an actuating motor, performing net positive mechanical work with two main power phases: 1) negative mechanical power absorption at weight acceptance wherein the product of the plantarflexor moment and dorsiflexor velocity is negative; and 2) a significant positive mechanical power burst by the plantarflexors during push-off. Note that the hip mechanical power is relatively small and irregular. The phases of negative mechanical work during human walking present an opportunity to improve the actuator efficiency of robotic leg prostheses and exoskeletons by recycling some of the otherwise dissipated energy using series elastic actuators and/or backdriveable actuators with energy regeneration.

2.2.2 Series Elastic Actuators

Elasticity is a mechanical principle that promotes safe and efficient human-robot dynamic physical interactions, which is especially important for robotic leg prostheses and exoskeletons. One popular engineering design, pioneered by Pratt and Williamson at the Massachusetts Institute of Technology (USA) [64], is to connect a passive elastic element (e.g., mechanical spring) in series between the actuator and external load, known as a series elastic actuator. Compared to traditional rigid actuators used in robotics, series elastic actuators have lower output impedance; greater shock tolerance and efficiency during foot-ground impacts; higher backdrivability via lower reflected inertia; and the ability to store and return elastic energy during negative and positive mechanical work periods, respectively [64]. Energy recycling via series compliance can improve the actuator efficiency by reducing the peak power and energy consumption of the electric motor provided that the stiffness of the elastic element is optimally designed (i.e., the spring-mass system dynamics matches the external load, thus requiring only a reactionary torque by the motor) [15]. This actuator design is bioinspired such that the elastic element stores and returns mechanical

energy similar to human muscle-tendon units as characterized by Hill muscle models with both active contractile and series elastic elements [59].

As series elastic actuators can reduce the mechanical power and torque requirements of the electric motor, this can further improve locomotor efficiency by reducing the size and weight of the onboard motors and batteries. Energy efficiency in legged locomotion is typically measured using cost of transport (COT),

$$COT = \frac{E}{Mgd} \tag{1}$$

where E is the energy consumed by a system of mass (M) to travel distance (d). For example, the Cassie bipedal robot, designed based on passive dynamics and series elastic actuation, achieved a cost of transport of ~0.7 such that the 30 kg robot consumes 200 W of electrical power while

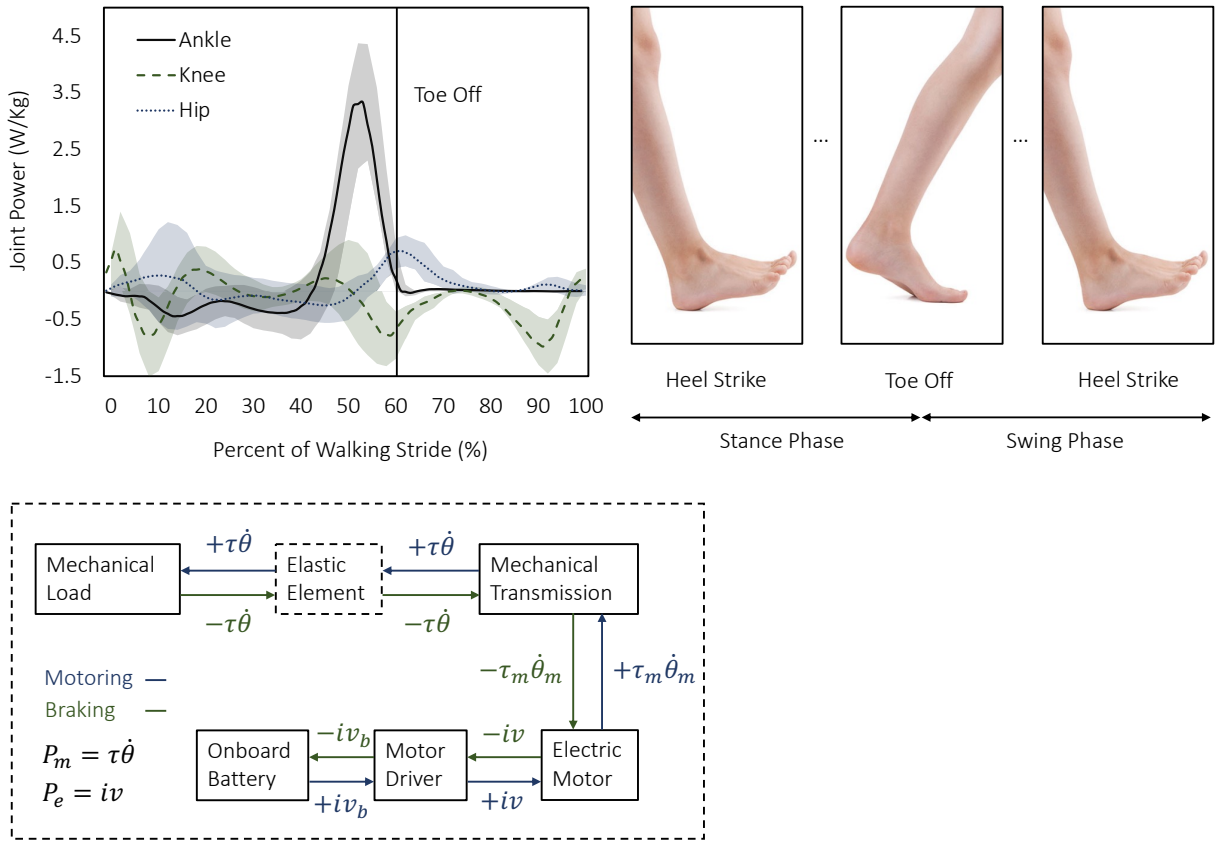

Figure 2.4. The average hip, knee, and ankle joint mechanical power (W/kg) per stride in healthy young adults (n=19) walking at natural cadence (110 ±8 steps/min at 1.44 m/s) and normalized to total body mass (top left). The positive and negative values represent joint mechanical power generation and absorption, respectively. Data were taken from Winter [54], the trajectories of which begin and end with heel-strike (top right). These locomotor energetics have implications on energy-efficient actuation of robotic leg prostheses and exoskeletons during motoring and braking operations (bottom left); the nomenclature are described in the text.

walking at 1 m/s [15]. In comparison, humans have a cost of transport of around 0.2. The hydraulically-actuated Big Dog quadrupedal robot by Boston Dynamics has a cost of transport of ~15 [15].

For applications in wearable robotic devices, Collins' group at Stanford University (USA) recently modelled and optimally controlled an energy-recycling actuator using an electroadhesive clutch and spring arranged in parallel with the electric motor [65]. Their simulation results showed that including parallel elasticity in their actuator design reduced the electrical power consumption by ~57%. In another example, Gregg's research lab at the University of Michigan (USA) used non-parametric convex optimization to optimize the stiffness of the elastic element to minimize peak power and energy consumption for arbitrary reference trajectories while satisfying actuator constraints [66]–[69]. Adding the optimized spring to their robotic ankle prosthesis reduced the peak power and energy consumption during walking from 450W to 132W and from 33 J to 25 J per stride, respectively [69].

The Biomechatronics group at MIT developed several robotic knee prostheses with series elasticity [70]–[73]. One prototype included a continuously-variable transmission between the motor and elastic element to operate the motor at optimal torque-speed regimes with highest efficiency by continuously varying the gear ratio [73]. Another prototyped included a clutchable series elastic actuator, whereby an electromagnetic clutch in parallel with the series elastic actuator supplied a reactionary torque when the task dynamics were elastically conservative and mechanical energy was recycled by the spring element [70]. The clutchable series elastic actuator consumed ~70% less electrical energy during walking than a series elastic actuator without the clutch mechanism [70]. Despite the aforementioned benefits, actuators with series elasticity tend to have lower output torque; increased weight and architecture complexity due to the added physical spring; and limited force and torque control bandwidth [5].

2.2.3 Regenerative Actuators

In recent years, torque-dense motors with low-ratio transmissions, known as (quasi) direct-drive actuators, have been used to achieve low mechanical impedance and high backdrivability and energy efficiency [74]. The application of low transmission ratios in legged and wearable robotic devices has largely been driven by advances in high torque-density brushless DC motors from the drone industry. This actuator design generates high output torque by increasing the motor torque density (torque per unit mass) rather than the transmission ratio, therein reducing the effects of high gearing (i.e., increased weight, backlash, and reflected inertia and damping of the actuator, which scale with the transmission ratio squared) [12]. Gears also have torque-dependent Coulomb friction that further increase mechanical impedance and reduce backdrivability and efficiency [11]; high external loads are thus needed to overcome the impedance to backdrive the actuator. These characteristics of high gearing systems can impede dynamic physical interactions between the

human and device and between the device and environment, which could especially encumber those with partial motor control function (e.g., elderly and/or individuals with osteoarthritis or poststroke) who could potentially benefit from the ability to backdrive the joints and actively participate in locomotion.

Compared to traditional rigid actuators used in robotics, backdriveable actuators with low mechanical impedance have many benefits for control and efficiency, including: 1) free-swinging dynamic leg motion, which can simplify the control during swing phase and allow for more natural, energy-efficient locomotion; 2) compliant foot-ground impacts; 3) negligible unmodeled actuator dynamics, which further simplifies the control; 4) intrinsic backdriveability comparable to series elastic actuators without their design and manufacturing complexities and low bandwidth; and 5) energy regeneration during negative mechanical work [74]. Energy regeneration is the process of converting some of the otherwise dissipated energy during negative mechanical work into electrical energy by backdriving the electromagnetic actuator. In other words, when backdriven by an external load, the motor can provide a braking torque to decelerate the load (e.g., motion control during swing phase) while simultaneously generating electricity [32]. This is analogous to regenerative braking in electric and hybrid electric vehicles. Assuming a regenerative motor driver, the regenerated energy could be used for battery recharging and/or transferred to other joints to support positive power generation. Such energy transmission and transduction is mathematically described by the laws of thermodynamics.

The autonomous quadrupedal robot known as the MIT Cheetah was one of the first legged robots to use backdriveable actuators with energy regeneration [74]–[78]. To minimize inefficiencies in their system (i.e., Joule heating losses in the electromagnetic motor; friction losses in the transmission; and interaction losses between the robot and environment via inelastic foot-ground impacts), the following design principles were implemented: 1) high torque-density motors with minimal gearing; 2) regenerative motor drivers; and 3) low leg mass and inertia. The motor torque density was increased by increasing the gap radius, which is the radius of the gap between the stator windings and permanent magnets on the rotor, assuming a constant mass. The low-ratio transmission (6:1) allowed for efficient bidirectional power flow between the motor and end effector during forward operation (i.e., power generation) and backdrive operation (i.e., power regeneration). The forward and backdrive directional efficiencies of the transmission were 98% and 96%, respectively, the differences of which were attributed to asymmetric friction and viscous damping losses [75]. The actuator backdrive efficiency was ~63% [74]. Exploiting these design principles, the MIT Cheetah achieved a cost of transport of ~0.5 such that the 33-kg robot can run at 6 m/s while consuming 973 W of electrical power [74].

Taking inspiration from the MIT Cheetah, Gregg's group [79]–[85] applied these actuator design principles to robotic leg prostheses and exoskeletons to achieve a low impedance, high backdriveable interface between the human and device during locomotion. They designed

pancake-style brushless DC motors with encapsulated windings for high torque-density. The large diameter of the motor allowed for a low-ratio transmission (7:1) to be designed directly inside the stator for a high package factor, which reduced the actuator size and weight. Benchtop and human walking experiments with a robotic leg prosthesis and exoskeleton showed that their actuator can generate 20-24 Nm of peak output torque and 1-3 Nm of backdrive torque [79]–[85], therein providing a high torque output during stance phase and low backdrive torque during swing phase. Their backdriveable actuators also allowed for energy regeneration and sharing between joints for improved locomotor efficiency [83]. To date, these wearable robotic devices are some of the few to demonstrate both power generation and regeneration during walking.

The knee exoskeleton by Donelan and colleagues at Simon Fraser University (Canada) [86]–[90] was designed to convert human biomechanical power into electrical power without requiring significant metabolic effort. Their system included a passive one-way clutch connected to spur gears with a 113:1 transmission ratio; a brushless DC generator; and a real-time control system to engage energy harvesting during late swing knee extension. Backdriving the motor during this period assisted the human muscles in performing negative mechanical work to decelerate the swinging leg prior to heel strike, therein minimizing the metabolic cost of operating the muscles as biological brakes, while concurrently generating electricity. The mechanical-to-electrical power conversion efficiency of the actuator was 63% [89]. The system performance was evaluated by the cost of harvesting (COH), which is the additional metabolic effort needed to generate electrical power $\left\{COH = \frac{\Delta metabolic\ power}{\Delta electrical\ power}\right\}$. When walking at 1.5 m/s, users were able to generate 4.8 ± 0.8 W of electricity with a 5 ± 21 W increase in metabolic power consumption compared to walking with the device but not generating electricity, therein yielding a cost of harvesting of 0.7 ± 4.4 [89]. This biomechanical energy harvesting knee exoskeleton could be worn on the unaffected limb of persons with unilateral impairments to recharge a robotic prosthesis or exoskeleton worn on the contralateral affected limb.

More recently, a collaborative group at Cleveland State University (USA) published a series of papers on modeling, optimization, and control of robotic and prosthetic systems with energy regeneration and storage [91]–[105]. Generally speaking, their multibody system models included a mechanical transmission coupled to a brushless DC motor; a regenerative motor driver; and an ultracapacitor for energy storage. Biogeography-based optimization, which is an evolutionary algorithm, was used to find the optimal system design and control parameters that maximized both energy regeneration and reference tracking motion control. A Pareto front was used to quantitatively evaluate the optimal trade-off between the two conflicting objective functions such that a higher impedance system tends to yield more accurate motion tracking but less energy regeneration. For example, their multi-objective optimization in [98] produced 0.0157 rad of RMS tracking error relative to reference joint kinematics while regenerating ~53 J of electrical energy

over 5-second computer simulations of walking with a semi-powered knee prosthesis. The energy regeneration efficiency was ~30%.

Although regenerative actuators can improve the energy efficiency and reduce the power consumption of robotic leg prostheses and exoskeletons, which can extend the battery-powered operating durations or decrease the weight of the onboard batteries, most previous studies have focused on steady-state level-ground walking [69], [70], [86]–[89], [91], [94], [97]–[100], [100]–[102], [104], [106]–[111]. However, real-world community mobility involves many short-duration walking bouts with low numbers of sequential steps (e.g., ~40% of walking bouts are less than 12 consecutive steps) [112]. Furthermore, targeted users of these wearable robotic devices (i.e., older adults and/or persons with physical disabilities) tend to walk slower and take fewer steps per day. For instance, research has shown that self-selected walking speed and daily step count decrease by 24% from 25 to 75 years age and by 75% from 60 to 85 years age, respectively [2]. These slower walking speeds have implications on energy regeneration performance since faster walking tends to generate more electricity and more efficiently [83], [88]–[90], [104], [106], [107], [110], [113]. These limitations in energy regeneration only from steady-state level-ground walking motivated the thesis research in Chapter 4. This review of robotic leg prostheses and exoskeletons with regenerative actuators was published in the ASME Journal of Mechanisms and Robotics [32].

2.3 Chapter Summary

In summary, this chapter reviewed the latest advances in automated control systems and energy-efficient actuators for robotic leg prostheses and exoskeletons.

Section 2.1 reviewed the development of locomotion mode recognition systems used for automated high-level control and decision making. Fusing data from mechanical and/or inertial sensors with surface EMG, known as neuromuscular-mechanical data fusion, can improve the locomotion mode recognition accuracies and decision times compared to implementing either system individually. However, neuromuscular-mechanical data are user-dependent and have limited prediction horizons such that the data can only characterize the current state of the user and/or device. Taking inspiration from the human vision-locomotor control system, wearable camera systems can perceive the oncoming walking environment prior to physical interactions, therein allowing for more accurate and robust high-level control decisions. For image classification of the walking environments, researchers have used statistical pattern recognition and machine learning algorithms like support vector machines, which require manual feature engineering that can be time-consuming and suboptimal. In contrast, deep learning replaces hand-designed features with multilayer networks that can automatically and efficiently learn the optimal image features from training data. Convolutional neural networks also tend to outperform support vector machines for image classification. These design and performance benefits of deep learning over traditional image classification algorithms motivated the thesis research in Chapter 3.

Section 2.2 reviewed the state-of-the-art in efficient actuators that exploit the energetics of legged locomotion and/or other energy storage and return mechanisms. The phases of negative mechanical work during human walking present an opportunity to improve efficiency by recycling energy using series elastic actuators and/or backdriveable actuators with energy regeneration. While series elastic actuators can reduce the peak power and energy consumption of the motor, these actuators tend to have increased weight and architecture complexity due to the added mechanical spring, and limited force and torque control bandwidth. Backdriveable actuators with low impedance transmissions can likewise facilitate dynamic locomotion, in addition to energy regeneration. Energy regeneration converts some of the otherwise dissipated energy during negative mechanical work into electrical energy, which can extend the battery-powered operating durations or decrease the weight of the onboard batteries. However, previous studies of robotic leg prostheses and exoskeletons with regenerative actuators have focused on steady-state level-ground walking. Given that targeted users of these wearable robotic devices tend to walk slower and take fewer steps per day, and that energy regeneration and efficiency both positively relate to walking speed, regenerating energy from steady-state level-ground walking is relatively limited, which motivated the thesis research in Chapter 4.

3. ENVIRONMENT RECOGNITION

In this chapter, a multi-generation environment sensing and classification system powered by computer vision and deep learning was developed to improve the automated high-level control and decision making of robotic leg prostheses and exoskeletons. The environment recognition system was designed to predict the oncoming walking environments prior to physical interaction, therein allowing for more accurate and robust locomotion mode transitions, which are currently limited to data from mechanical, inertial, and/or neuromuscular sensors.

3.1 Preliminary System Design

3.1.1 Introduction

As reviewed in Chapter 2, previous environment recognition systems for robotic leg prostheses and exoskeletons have used statistical pattern recognition and machine learning algorithms like support vector machines for image classification, which require manual feature engineering that can be time-consuming and suboptimal (see Table 2.2). In contrast, deep learning replaces hand-designed features with multilayer networks that can automatically and efficiently learn the optimal image features from training data. Convolutional neural networks also tend to outperform support vector machines for image classification [55]. Deep learning has become pervasive in computer vision ever since the AlexNet network [114], developed by Hinton's research lab at the University of Toronto (Canada), won the 2012 ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [115] with a top-1 classification accuracy of ~63%.

Motivated by these design and performance benefits, the objective of this research study was to develop a preliminary environment recognition system powered by deep learning to evaluate the feasibility of using convolutional neural networks for this novel computer vision application. Images of indoor and outdoor real-world walking environments were collected around the University of Waterloo campus using a lightweight wearable camera and a deep convolutional neural network was designed and trained to predict the oncoming walking environments, specifically level-ground terrain, incline stairs, and decline stairs. Given that the accuracy of a locomotion mode recognition system used for high-level control is especially important when interacting with stair environments, from a safety critical perspective, this preliminary system design focused on stair recognition.

3.1.2 Image Dataset of Walking Environments

One participant was instrumented with a chest mounted RGB camera system (GoPro Hero4) (see Figure 3.1). Unlike lower-limb systems [31], [35]–[37], [43]–[45], [49], [50], [116], chest-mounting can provide more stable video recording and allow users to wear pants and dresses without obstructing the sampled field-of-view. Whereas most environment sensing and classification

systems have been limited to controlled laboratory environments and/or prearranged walking circuits [23], [24], [26], [29], [34], [35], [38], [39], [46]–[50], [54], [116], the participant in this study walked around the University of Waterloo while collecting RGB images of unknown outdoor and indoor real-world environments with variable lighting, occlusions, signal noise, and intraclass variations. Data were collected at various times throughout the day to account for different lighting conditions. The visual field-of-view was ~3 m ahead of the participant. Images were sampled at 60 Hz with a 1280×720 resolution and stored on a 64-GB microSD memory card. The wearable camera system weighs ~0.3 kg and includes an onboard rechargeable lithium-ion battery. Approximately 10 hours of video were recorded across ten 1-hour walking sessions, amounting to ~2 million images. Ethical review and approval were not required for this research study in accordance with the University of Waterloo Office of Research Ethics.

Since there were minimal differences between consecutive images sampled at 60 Hz, the dataset was downsampled to 1 frame/second to train the convolutional neural network. However, for online environment-adaptive control of robotic leg prostheses and exoskeletons, higher sampling rates would be advantageous for more accurate and robust automated locomotion mode transitions. Images were cropped to an aspect ratio of 1:1 and resized to 224x224 using bilinear interpolation. 34,254 images were processed and manually labelled into one of three environment classes, including 27,030 images of level-ground terrain; 3,943 images of incline stairs; and 3,281 images of decline stairs. Manually labelling the different walking environments was relatively subjective, especially near class transitions. For transitions from stairs to level-ground terrain, the images were labelled “stairs” whenever the stairs were approximately within the visual field-of-view. For transitions from level-ground terrain to stairs, the images were again labelled “stairs” once the first step of the staircase became visible within the field-of-view. Images were labelled by one designated researcher for consistency.

3.1.3 Convolutional Neural Network

A 10-layer convolutional neural network was designed and trained for environment classification (Figure 3.1) using TensorFlow 1.1 [117]. The architecture included multiple stacked convolutional and pooling layers with decreasing spatial resolutions and increasing number of feature maps. Starting with an input RGB image, the convolutional layers performed convolution operations (i.e., dot products) between the inputs and convolutional filters; the network mainly used 3x3 filters. The first few layers extract relatively general features, like edges, while deeper layers learn more abstract, problem-dependent features. The feature maps were then passed through a rectified linear unit (ReLU) nonlinear activation function. ReLu nonlinearities tend to yield faster training times compared to other activation functions [55]. The pooling layers spatially downsampled the feature maps to reduce the computational effort by aggregating neighboring elements into a single feature map using maximum values; spatial downsampling was also performed using convolutional layers with strides of two. The CNN architecture concluded with a fully connected

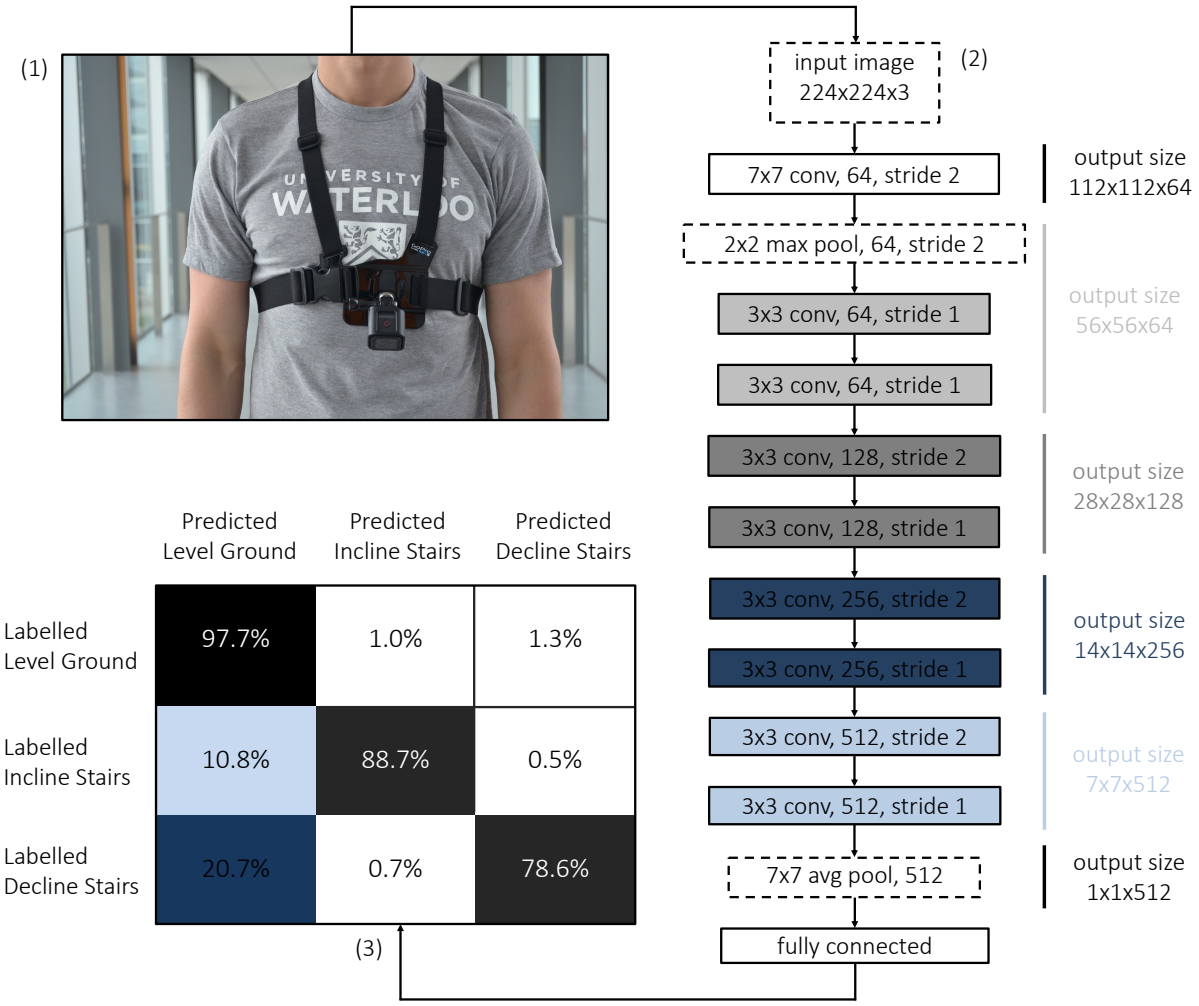

Figure 3.1. Schematic of the preliminary environment recognition system powered by computer vision and deep learning. Images of real-world walking environments were collected using a lightweight wearable camera (1) and classified using a 10-layer convolutional neural network (2). The classification predictions during inference for each environment class are displayed using a multiclass confusion matrix (3).

layer and a softmax loss function, which estimates the probability distribution (i.e., the predicted scores) of each labelled class. The neural network was trained using supervised learning such that the differences between the predicted and labelled class scores were computed and the learnable network parameters (i.e., weights) were optimized to minimize the loss function through backpropagation and stochastic gradient descent.

Five-fold cross-validation was used during training such that the dataset was split into five individual folds and each fold was trialed for validation. Each data-collection session was thus used for both training and validation. Since there were significant class imbalances in the dataset,

images of level-ground terrain and stairs were undersampled and oversampled during training, respectively, according to the class priors from each fold to maintain uniform distributions. Input images to the neural network were normalized by subtracting the average pixel intensities from each fold. The learnable weights were initialized with Xavier initialization [118]. The neural network was trained for 2000 iterations using a batch size of 256 (~19 epochs); the Adam optimizer [119]; and an initial learning rate of 0.001, which was halved after 1000 training iterations. Images were also randomly flipped horizontally during training at a rate of 0.5 to increase stochasticity and prevent overfitting. Dropout regularization was experimented with but showed no additional performance benefit. After training, the CNN performed inference to evaluate the generalizability of the learned parameters. Training and inference were both performed on an NVIDIA TITAN Xp GPU. Training took ~20 minutes per fold. The computational cost of the convolutional neural network was 1.75-G floating point operations (FLOPs).

3.1.4 Classification Results

Figure 3.1 shows the multiclass confusion matrix, which illustrates the classification performance of the convolutional neural network during inference. The horizontal and vertical axes are the predicted and labelled classes, respectively. The diagonal elements are the image classification accuracies (%) for each environment class, as known true positives, and the nondiagonal elements are the misclassification percentages. The environment recognition system achieved 94.85% image classification accuracy during inference, that being the percentage of true positives (32,491 images) out of the total number of labelled images (34,254 images). The convolutional neural network most accurately predicted level-ground terrain, followed by incline stairs, and decline stairs. These results could be attributed to the class imbalances in the dataset such that there were significantly more images of level-ground terrain compared to other classes. There were relatively few misclassifications between incline stairs and decline stairs.

Figure 3.2 shows several examples of failure cases whereby the convolutional neural network incorrectly predicted the walking environment. The images in the top, middle, and bottom rows were misclassified as incline stairs, level-ground, and decline stairs, respectively. The misclassifications in the top row (i.e., images of level-ground terrain misclassified as incline stairs) were particularly interesting given that these misclassifications seem to challenge even the human visual system. The misclassifications in the middle row (i.e., images of decline stairs misclassified as level-ground terrain) demonstrate the importance of classification accuracy for safe locomotor control such that inaccurate control decisions, due to errors in the environment prediction, could result in the user falling down a flight of stairs. The misclassifications in the bottom row highlight a limitation of the labelling architecture such that the convolutional neural network was forced to classify images with “random” objects and/or environments into one of the three predefined classes (i.e., either level-ground terrain, incline stairs, or decline stairs).

In summary, images of indoor and outdoor real-world walking environments around the University of Waterloo campus were collected using a wearable camera and a deep convolutional neural network was designed and trained for environment recognition, achieving 94.85% image classification accuracy. This preliminary environment sensing and classification system was published in the 2019 IEEE International Conference on Rehabilitation Robotics (ICORR) [40]. Since then, research in environment recognition for robotic leg prostheses and exoskeletons has almost entirely shifted towards using deep learning for image classification [31], [38], [42], [43], [48]–[51], [116]. For example, the convolutional neural network developed by Simon's group at Cleveland State University (USA) [38] achieved ~99% image classification accuracy across three environment classes using transfer learning of pretrained weights. Fu's research lab at the Southern University of Science and Technology (China) published a series of papers on using deep learning for 3D environment sensing and classification for prosthetic leg control [31], [48]–[50]. However, these newly-implemented convolutional neural networks for environment recognition (Table 3.1 and Figure 3.3) have each been trained and tested on different image datasets, therein preventing direct comparisons between CNN architectures from different researchers, which motivated the following research in Section 3.2.

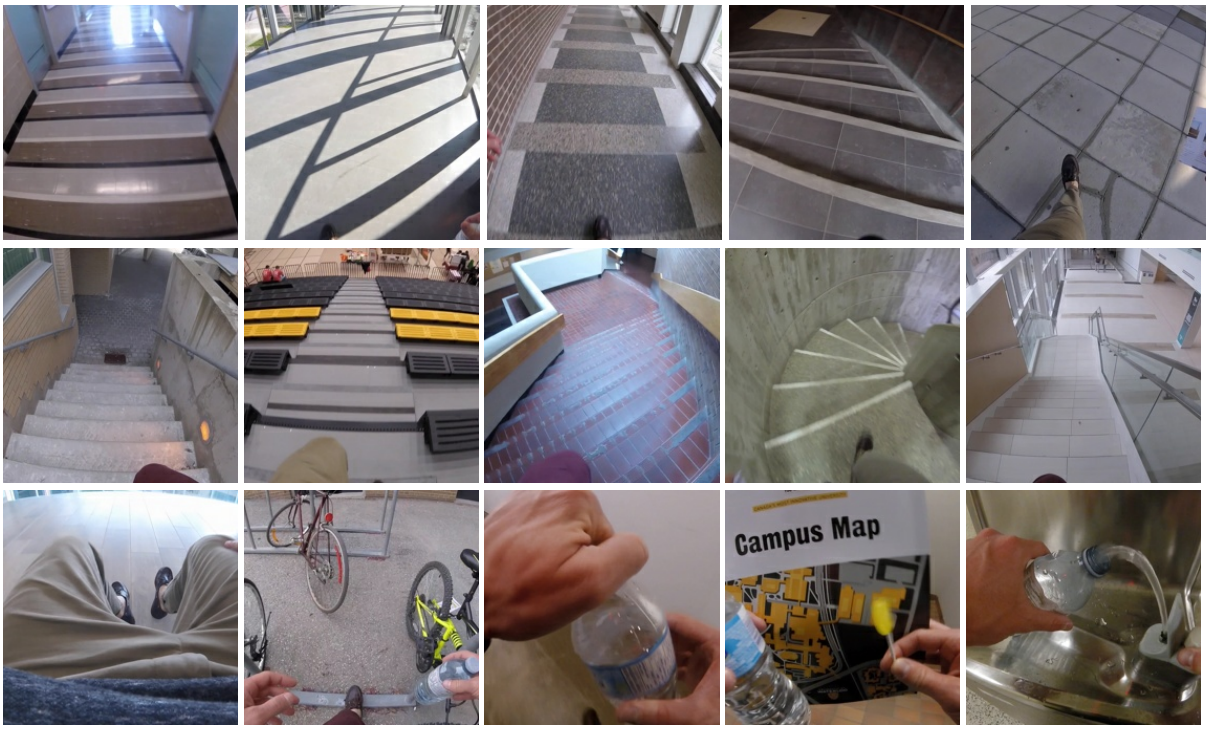

Figure 3.2. Examples of failure cases whereby the convolutional neural network incorrectly predicted the walking environment. Images in the top, middle, and bottom rows were misclassified as incline stairs, level-ground, and decline stairs, respectively.

Table 3.1. Environment recognition systems that used convolutional neural networks for image classification of walking environments. Note that each neural network was trained and tested on different image datasets (Table 2.1). The computing hardware were mostly developed and manufactured by NVIDIA. The number of operations is expressed in multiply-accumulates. The inference times are reported per image.

Reference	Operations (billions)	Parameters (millions)	Computing Devices	Test Accuracy (%)	Inference Time (ms)
Khademi and Simon [38]	7.7	27	Titan X	99.6	50
Laschowski et al. [40]	1.2850	4.73	TITAN Xp	94.9	0.9
Nouredanesh et al. [53]	0.0020	0.04	Intel Core i7 (3.60GHz)	92.0	N/A
Novo-Torres et al. [42]	0.0011	1.13	Geforce GTX 965M	90.0	5.5
Zhang et al. [50]	0.0130	0.22	GeForce GTX 1050 Ti	96.8	3.1
Zhang et al. [49]	0.0130	0.22	Quadro P400	98.9	3.0
Zhang et al. [48]	0.0215	0.05	GeForce GTX 1050 Ti	99.0	2.0
Zhang et al. [31]	0.0130	0.22	Quadro P400	96.0	3.0
Zhong et al. [43]†	0.0544	2.20	Jetson TX2	95.4	12.7
Zhong et al. [51]λ	0.0544	2.20	Jetson TX2	94.0	13.1

†used MobileNetV2 and λused MobileNet for the feature extraction and both used a Bayesian neural network (i.e., gated recurrent unit) for the environment classifications.

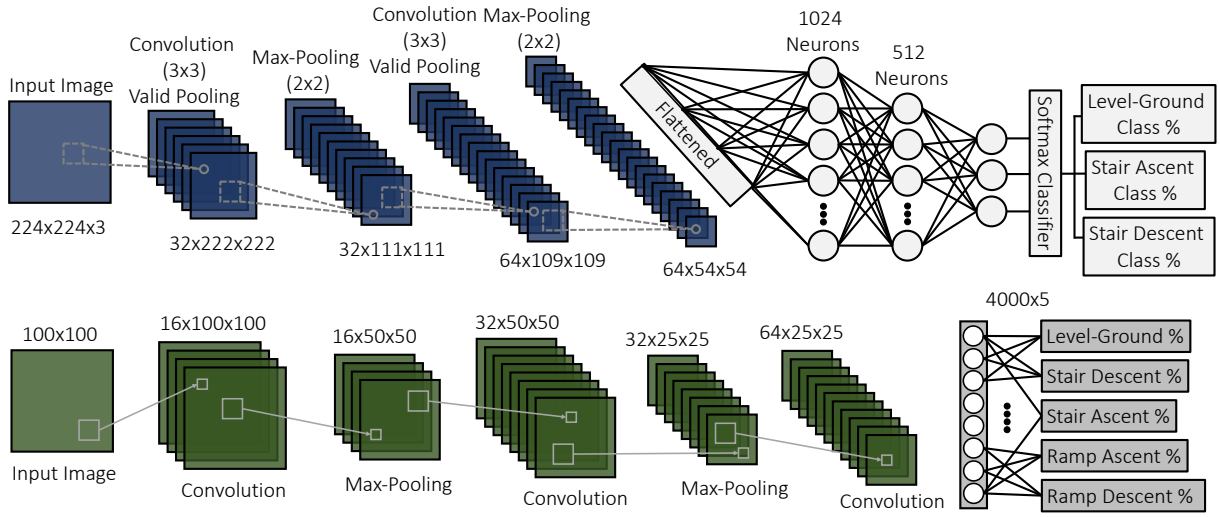

Figure 3.3. Examples of convolutional neural networks used for image classification of walking environments. The top and bottom schematics were adapted from Drs. Dan Simon (Cleveland State University, USA) and Chenglong Fu (Southern University of Science and Technology, China), respectively.

3.2 Second-Generation System

3.2.1 Introduction

Advances in deep learning and computer vision have recently allowed researchers to develop environment recognition systems to supplement the locomotion mode recognition systems used for automated high-level control and decision making of robotic leg prostheses and exoskeletons. However, small-scale and private image datasets have prevented the widespread development and comparison of convolutional neural networks for environment classification (Table 3.1). To date, researchers have each individually collected training data to develop their image classification algorithms. These repetitive measurements are time-consuming and inefficient, and individual private datasets have prevented direct comparisons between convolutional neural networks from different researchers.

Taking inspiration from ImageNet [120], here the "ExoNet" database was developed – the largest and most diverse open-source dataset of wearable camera images (i.e., egocentric perception) of real-world walking environments. Over a dozen state-of-the-art deep convolutional neural networks were trained and tested on ExoNet for image classification and automatic feature engineering. The benchmarked CNN architectures and their environment classification predictions were quantitatively evaluated and compared using an operational metric called "NetScore", which balances the image classification accuracy with the architectural and computational complexities (i.e., important for onboard real-time inference with mobile computing devices). This second-generation environment sensing and classification system provides a large-scale benchmark and reference for future research in environment-adaptive locomotor control.

3.2.2 Large-Scale Image Dataset

One participant, without wearing and/or using an assistive device (e.g., an exoskeleton, walker, or crutches), was instrumented with a lightweight smartphone camera system (iPhone XS Max) (see Figure 3.4). The chest-mount height was ~1.3 m from the ground when the participant stood upright. The smartphone has two 12-megapixel RGB rear-facing cameras and one 7-megapixel front-facing camera. The front and rear cameras provide 1920×1080 and 1280×720 video recording at 30 Hz, respectively. The smartphone weighs 0.21 kg and has an onboard rechargeable lithium-ion battery; 512-GB of memory storage; and a 64-bit ARM-based integrated circuit (Apple A12 Bionic) with a six-core CPU and four-core GPU. These hardware specifications can theoretically support onboard deep learning inference for real-time environment classification. The relatively lightweight and unobtrusive nature of the wearable camera system allowed for unimpeded human locomotion. Similar to Section 3.1, ethical review and approval were not required for this research study in accordance with the University of Waterloo Office of Research Ethics.

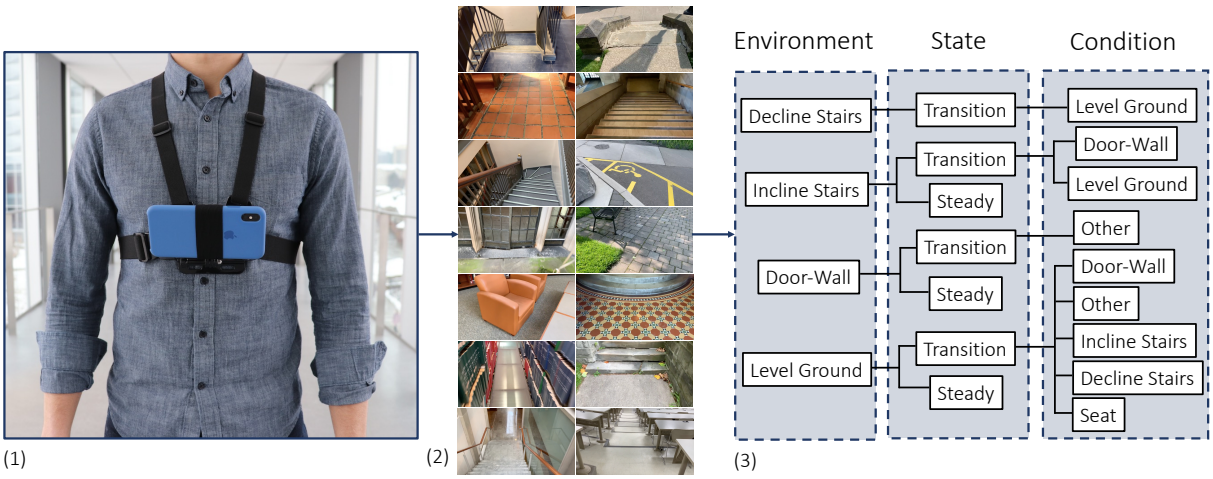

Figure 3.4. Development of the "ExoNet" database, including (1) a photograph of the wearable camera system used for large-scale data collection; (2) examples of the high-resolution RGB images of walking environments; and (3) a schematic of the novel hierarchical labelling architecture with 12 classes.

The subject walked around unknown outdoor and indoor real-world environments while collecting images with occlusions and intraclass variations (Figure 3.5). Data were collected at various times throughout the day to include different lighting conditions. Inspired by human gaze fixation during walking [121], the visual field-of-view was 1-5 m ahead of the participant, thereby showing the oncoming walking environment rather than the ground directly underneath the subject's feet. This operating range allowed for detecting obstacles and terrain changes within several walking strides; research has shown that lower-limb amputees tend to allocate higher visual attention to oncoming terrain changes compared to able-bodied individuals [121]. The camera's pitch angle slightly differed between data collection sessions. Images were sampled at 30 Hz with 1280×720 resolution. Over 52 hours of video were recorded, amounting to ~5.6 million images. The same environment was never sampled twice to maximize diversity in the dataset. Images were collected throughout the summer, fall, and winter seasons to capture different weathered surfaces like snow, grass, and multicolored leaves. The image database, which was named "ExoNet", was uploaded to IEEE DataPort and is now publicly available for download at https://ieee-dataport.org/open-access/exonet-database-wearable-camera-images-human-locomotion-environments [41]. The file size of the uncompressed videos is ~140 GB.

Given the subjects' self-selected walking speed, there were relatively minimal differences between consecutive images sampled at 30 Hz. The images were thus downsampled and labelled at 5 frames per second to minimize the demands of manual annotation and increase the diversity in image appearances. Similar to the ImageNet dataset [120], the ExoNet database was human-annotated using a hierarchical labelling architecture. The images were mainly labelled according to common high-level locomotion modes of robotic leg prostheses and exoskeletons, rather than

a purely computer vision perspective. For instance, images of level-ground terrain showing either pavement or grass were not differentiated since both surface textures would be assigned the same high-level locomotion mode (i.e., level-ground walking). However, with advances in control system designs, future research may want to consider the image classification of different walking surface textures.

Approximately 923,000 images were manually annotated using a novel hierarchical labelling architecture with 12 classes (Figure 3.4). The dataset included: 31,628 images of "incline stairs transition wall/door" (I-T-W); 11,040 images of "incline stairs transition level-ground" (I-T-L); 17,358 images of "incline stairs steady" (I-S); 28,677 images of "decline stairs transition level-ground" (D-T-L); 19,150 images of "wall/door transition other" (W-T-O); 36,710 images of "wall/door steady" (W-S); 379,199 images of "level-ground transition wall/door" (L-T-W); 153,263 images of "level-ground transition other" (L-T-O); 26,067 images of "level-ground transition incline stairs" (L-T-I); 22,607 images of "level-ground transition decline stairs" (L-T-D); 119,515 images of "level-ground transition seat" (L-T-E); and 77,576 images of "level-ground steady" (L-S) (see Figure 3.6). These class labels were developed and assigned post hoc to encompass the different walking environments from the data collection. In comparison, Novo-Torres et al. [42] simultaneously sampled and labelled their images online using a portable keyboard during data collection. Similar to previous research [24], [26], [34], an "other" class was used to maintain the image classification performance when unlabeled environments and/or objects like pedestrians, cars, and bicycles were observable. Having not included an "other" class in the preliminary environment recognition system in Section 3.1, the convolutional neural network was required to classify random objects and/or environments, such as those shown in Figure 3.2, into one of the three predefined classes, which likely affected the overall image classification accuracy.

Taking inspiration from [23], [24], [26], [28], [38], [54], the ExoNet hierarchical labelling architecture included both steady (S) and transition (T) states (Figure 3.7). A steady state describes an environment where an exoskeleton or prosthesis user would continue to perform the same locomotion mode (e.g., an image showing only level-ground terrain). In contrast, a transition state describes an environment where an exoskeleton or prosthesis high-level controller might switch between different locomotion modes (e.g., an image showing both level-ground terrain and incline stairs). Manually labelling these transition states was relatively subjective. For instance, an image showing level-ground terrain was labelled "level-ground transition incline stairs" (L-T-I) when an incline staircase was approximately within the visual field-of-view. Similar labelling principles were applied to transitions to other conditions. Although the ExoNet database was labelled by one designated researcher, consistently determining the exact video frame where an environment should switch between steady and transition states was challenging; Huang et al. [28] reported experiencing similar difficulties. The development of the ExoNet database was published in the Frontiers in Robotics and AI [41].

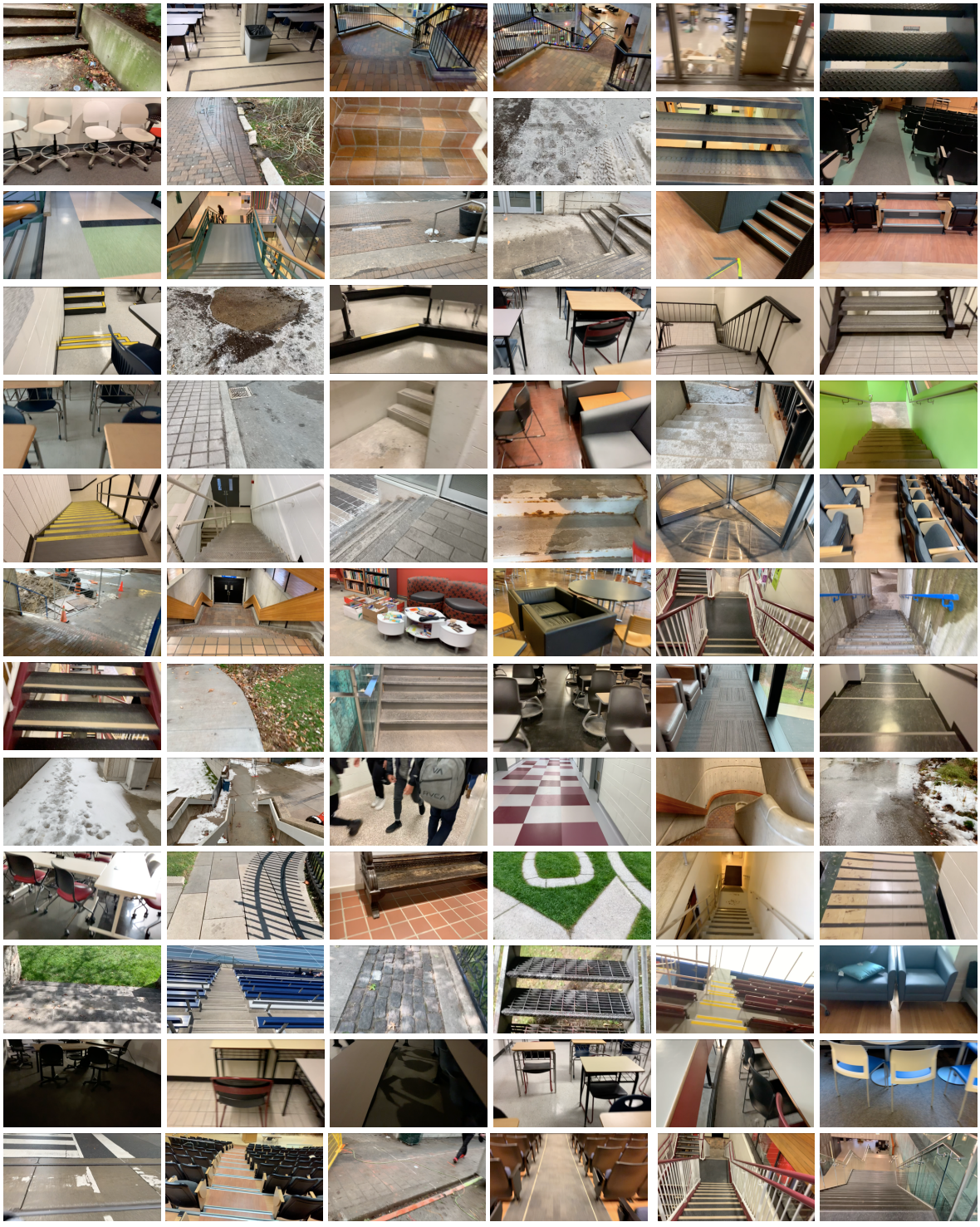

Figure 3.5. Examples of the wearable camera images of indoor and outdoor real-world walking environments in the ExoNet database. Images were collected at various times throughout the day and across different seasons (i.e., summer, fall, and winter).

Label	Number of Images	Percent of Dataset (%)
L-T-W	379,199	41.1
L-T-O	153,263	16.6
L-T-E	119,515	13.0
L-S	77,576	8.4
W-S	36,710	4.0
I-T-W	31,628	3.4
D-T-L	28,677	3.1
L-T-I	26,067	2.8
L-T-D	26,067	2.4
W-T-O	19,150	2.1
I-S	17,358	1.9
I-T-L	11,040	1.2
Total	922,790	100

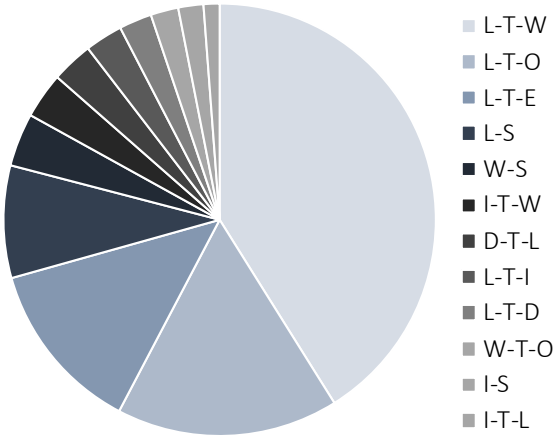

Figure 3.6. Schematic of the class distributions in the ExoNet database, the images of which were annotated using a novel hierarchical labelling architecture. A description of the class labels is provided in the text.

Note that the images were sampled with relatively high-resolution (1280×720) to provide a rich open-source dataset for the research community such that prospective users of ExoNet have the option to downscale the image sizes depending on their application. Furthermore, the labelling architecture was designed based on current high-level controllers. The most common locomotion modes programmed into robotic leg prostheses and exoskeletons include level-ground walking, stair ascent, stair descent, sit-to-stand, and stand-to-sit [4]. An environment recognition system should thus, at minimum, be designed to classify level-ground terrain, seats, incline stairs, decline stairs, and transitions between these different environments; accurate sensing and classification of transitions, especially when involving stairs, is important given that they present a safety-critical application. To reiterate, an image was labelled as a transition state when the field-of-view showed both the current walking environment (e.g., level-ground terrain) and another environment where an exoskeleton or prosthesis high-level controller might switch between locomotion modes (e.g., incline stairs), such as that illustrated in Figure 3.7. In this example, the beginning of the transition was defined as when the lip of the first step in the staircase was visible within the field-of-view. Although manually labelling the images at 5 frames per second, rather than the original sampling rate of 30 Hz, created greater diversity in the image appearances and thus helped with labelling the transitions, the beginning and end of the environment transitions was relatively subjective, which likely affects inter-rater reliability.

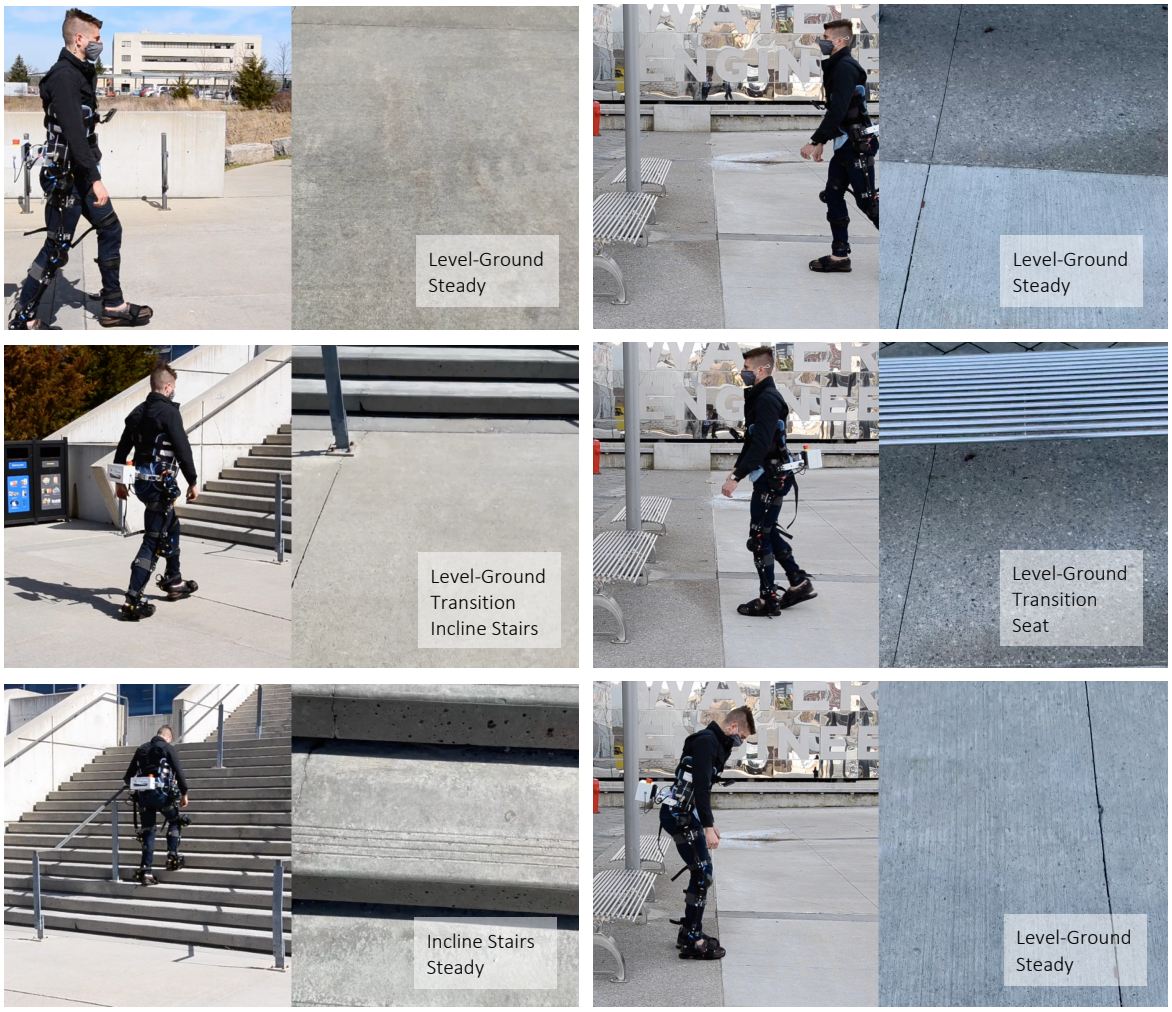

Figure 3.7. Examples of "steady" and "transition" states in the ExoNet hierarchical labelling architecture. The images in the top and bottom rows were labelled as steady states and those in middle row were labelled as transition states. For each column, the left images show the author walking with a robotic exoskeleton and the right images show the concurrent field-of-view of the wearable camera system (i.e., what the exoskeleton would see).

3.2.3 Deep Convolutional Neural Networks

TensorFlow 2.3 and the Keras functional API were used to build, train, and test over a dozen state-of-the-art deep convolutional neutral networks on the ExoNet database, including: EfficientNetB0 [122]; InceptionV3 [123]; MobileNet [124]; MobileNetV2 [125]; VGG16 and VGG19 [126]; Xception [127]; ResNet50, ResNet101, and ResNet152 [128]; and DenseNet121, DenseNet169, and DenseNet201 [129]; these artificial neural networks are biologically-inspired variants of the multilayer perceptron that resemble the human visual cortex. During data preprocessing, the images were cropped to an aspect ratio of 1:1 and downsampled to 256x256 using bilinear interpolation. Random crops of 224x224 were used as inputs to the neural networks; this method of data augmentation helps further increase the sample diversity. Unlike previous studies that used support vector machines that required hand-engineering [44], [45], convolutional neural networks can automatically and efficiently learn the optimal image features from training data. The final densely connected layer of each CNN architecture was modified by setting the number of output channels equal to the number of environment classes in the ExoNet database (n=12). A softmax loss function was used to predict the individual class scores of each labelled environment.

The labelled ExoNet images were split into training (89.5%), validation (3.5%), and testing (7%) sets, the proportions of which are consistent with ImageNet [120], which is of comparable size. Transfer learning of pretrained weights from ImageNet were experimented with but showed no additional performance benefit. Dropout regularization [130] was applied before the final dense layer to prevent overfitting during training such that the learnable weights were randomly dropped (i.e., activations set to zero) during each forward pass at a rate of 0.5. Images were also randomly flipped horizontally during training to increase stochasticity and promote generalization. Each CNN architecture was trained for 40 epochs using a batch size and initial learning rate of 128 and 0.001, respectively; these hyperparameters were experimentally tuned to maximize the performance on the validation set and averaged across trials (see Figure 3.8).

Different combinations of batch sizes of 32, 64, 128, and 256; epochs of 20, 40, and 60; dropout rates of 0, 0.2, 0.5; and initial learning rates of 0.01, 0.001, 0.0001, and 0.00001 were explored. The learning rate was reduced during training using a cosine weight decay schedule [131]. The sparse categorical cross-entropy loss between the labelled and predicted classes were calculated, and the Adam optimizer [119] was used to update the learnable weights and minimize the loss function by calculating backpropagated gradients using momentum and an adaptive learning rate. A single central crop of 224x224 was used during testing. Training and inference were both performed on a Tensor Processing Unit (TPU) version 3-8 by Google Cloud; these customized chips can allow for accelerated CNN computations (i.e., matrix multiplications and additions) compared to more traditional computing devices.

3.2.4 Classification Results

Table 3.2 summarizes the benchmarked CNN architectures (i.e., the number of parameters and computing operations) and their environment classification performances on the ExoNet database (i.e., prediction accuracies and inference times). The EfficientNetB0 network achieved the highest image classification accuracy (C_a) during inference (73.2% accuracy) on the ExoNet database, that being the percentage of true positives (47,265 images) out of the total number of images in the testing set (64,568 images) $\left\{C_a = \frac{True\ Positives}{Total\ Images} \times 100\%\right\}$. In contrast, the VGG19 architecture produced the least accurate predictions, with an image classification accuracy of 69.2%. The range of classification accuracies across the benchmarked CNN architectures was thus relatively small with a maximum arithmetic difference of only 4 percentage points. Relatively weak statistically correlations were observed between both the number of parameters (Pearson r = -0.3) and computing operations (Pearson r = -0.59) and the image classification accuracies on the ExoNet database across the benchmarked CNN architectures.

Although the VGG16 and VGG19 networks have the largest number of computations (i.e., 15.4 and 19.5 billion multiply-accumulates, respectively), they resulted in the fastest inference times (i.e., on average ~1.4 ms and ~1.6 ms per image). For comparison, the DenseNet201 network has 72.1% and 78% fewer computing operations than VGG16 and VGG19, respectively, but was 364% and 306% slower. Note that these inference times were calculated on the Cloud TPU using a batch size of eight. These performance trends concur with those reported by Ding et al. [132], who recently showed that 1) the number of computing operations does not explicitly reflect the actual inference speed; and 2) VGG-style architectures can run faster and more efficiently on CNN computing devices compared to more complicated architectures like DenseNets due to their relatively simple designs with basic convolutions and ReLU activations. However, the relative inference speeds between the benchmarked CNN architectures (i.e., their ordering from fastest to slowest) could theoretically differ across different computing devices given that some platforms are designed to accelerate certain operations better than others (e.g., cloud computing vs. those designed for mobile and embedded systems).

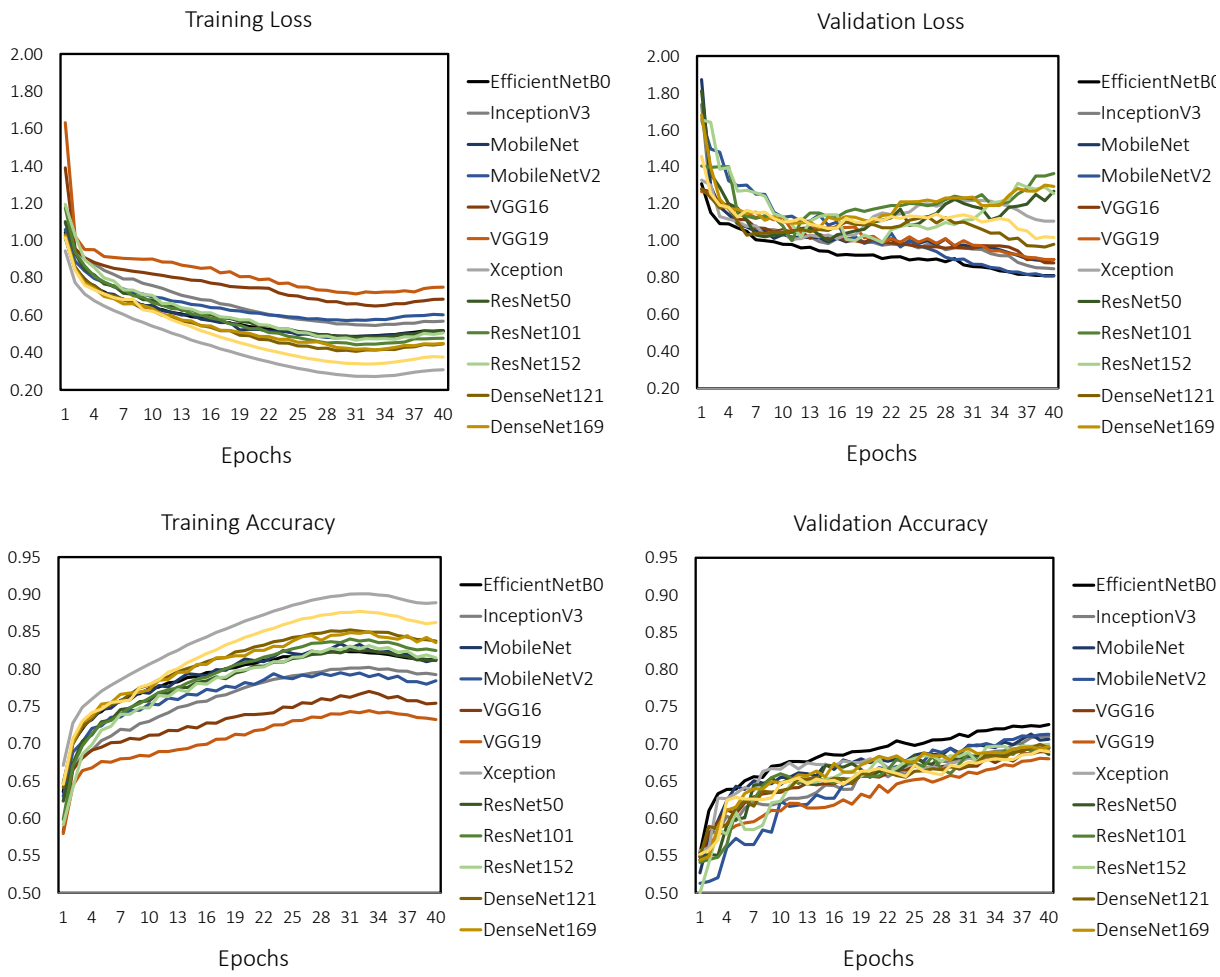

Figure 3.8. The loss and image classification accuracies during training and validation on the ExoNet database using state-of-the-art deep convolutional neural networks, including: EfficientNetB0, InceptionV3, MobileNet, MobileNetV2, VGG16, VGG19, Xception, ResNet50, ResNet101, ResNet152, DenseNet121, DenseNet169, and DenseNet201.

Table 3.2. The benchmarked CNN architectures and their environment classification performances during inference on the ExoNet database. The test accuracies, parameters, and computing operations are expressed in percentages (0-100%), millions of parameters (M), and billions of multiply-accumulates (GMACs), respectively. Training and inference were both performed on a Google Cloud TPU. The EfficientNetB0 network achieved the highest test accuracy; VGG16 the fastest inference time; and MobileNetV2 the best NetScore and least number of parameters and computing operations.

CNN Architecture	Operations (GMACs)	Parameters (M)	Test Accuracy (%)	Inference Time (ms)	NetScore (Ω)
EfficientNetB0	0.39	4.06	73.2	2.5	72.6
InceptionV3	2.84	21.83	71.9	4.1	56.3
MobileNet	0.57	3.24	71.1	1.6	71.4
MobileNetV2	0.30	2.27	72.9	2.2	76.2
VGG16	15.36	14.72	70.1	1.4	50.3
VGG19	19.52	20.03	69.2	1.6	47.7
Xception	4.55	20.89	70.4	2.3	54.1
ResNet50	3.86	23.61	69.5	2.5	54.1
ResNet101	7.58	42.68	70.1	4.2	48.7
ResNet152	11.29	58.40	71.6	5.6	46.0
DenseNet121	2.83	7.05	71.5	4.4	61.2
DenseNet169	3.36	12.66	70.7	5.7	57.7
DenseNet201	4.29	18.35	70.2	6.5	54.9

Table 3.3. The multiclass confusion matrix for EfficientNetB0 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	78.8	0.6	0.3	0.0	0.0	0.0	1.0	4.5	10.6	0.2	2.6	1.4
W-S	0.2	72.1	9.2	0.0	0.3	0.2	0.0	0.3	15.2	0.1	1.9	0.5
W-T-O	0.4	21.9	43.2	0.0	0.4	0.2	0.1	0.1	19.2	0.4	8.8	5.5
I-S	0.0	0.3	0.1	62.1	33.9	1.9	0.0	0.0	0.5	0.6	0.5	0.0
I-T-W	0.0	2.0	0.7	16.8	69.0	2.8	0.0	0.2	1.5	5.8	0.6	0.6
I-T-L	1.0	0.5	0.2	2.5	5.7	77.9	2.2	0.2	6.4	1.5	1.2	0.7
L-S	0.1	0.4	0.0	0.0	0.0	0.1	79.9	0.3	11.9	0.6	6.4	0.2
L-T-D	5.5	0.4	0.6	0.0	0.3	0.1	1.0	53.3	28.3	2.8	4.3	3.6
L-T-W	0.3	1.5	0.4	0.0	0.0	0.1	3.7	0.4	86.5	0.3	4.4	2.3
L-T-I	0.1	1.0	0.2	1.0	3.9	0.4	3.8	0.7	23.3	49.1	12.1	4.4
L-T-O	0.3	0.6	1.1	0.0	0.1	0.1	13.7	0.5	28.8	0.7	47.5	6.6
L-T-E	0.3	0.5	0.3	0.0	0.0	0.1	0.8	0.5	14.0	0.6	10.1	72.7

Table 3.4. The multiclass confusion matrix for InceptionV3 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	77.3	0.8	0.3	0.0	0.1	0.4	1.1	4.3	11.0	0.2	2.6	1.9
W-S	0.3	70.6	10.9	0.0	0.2	0.1	0.0	0.2	16.3	0.1	1.0	0.3
W-T-O	0.7	22.7	38.4	0.0	0.3	0.0	0.2	0.5	21.0	0.5	12.0	3.7
I-S	0.0	0.8	0.1	66.6	25.9	3.3	0.8	0.0	1.0	1.0	0.3	0.3
I-T-W	0.1	2.0	0.9	15.1	68.7	3.2	0.0	0.0	2.9	5.3	1.0	0.8
I-T-L	0.3	1.3	0.2	2.5	3.9	72.4	4.0	0.2	10.3	2.4	1.0	1.5
L-S	0.2	0.1	0.0	0.0	0.0	0.1	78.6	0.2	12.5	0.3	7.4	0.5
L-T-D	6.1	0.6	0.4	0.0	0.4	0.1	1.1	53.3	28.2	1.1	5.1	3.6
L-T-W	0.3	1.9	0.5	0.0	0.1	0.1	4.2	0.5	85.5	0.4	4.3	2.2
L-T-I	0.0	0.9	0.4	0.5	4.0	0.5	4.4	1.4	24.9	47.6	10.8	4.5
L-T-O	0.4	0.5	1.1	0.0	0.1	0.0	14.5	0.5	30.7	0.6	46.2	5.3
L-T-E	0.3	0.5	0.6	0.0	0.1	0.1	0.9	0.4	16.1	0.7	9.7	70.6

Table 3.5. The multiclass confusion matrix for MobileNet showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	77.7	0.7	0.5	0.0	0.1	0.2	1.0	3.6	10.8	0.2	3.6	1.6
W-S	0.1	73.0	9.8	0.0	0.3	0.0	0.1	0.3	14.2	0.2	1.5	0.6
W-T-O	2.7	24.2	38.4	0.0	0.1	0.1	0.2	0.4	19.8	0.5	10.0	3.4
I-S	0.3	0.4	0.8	59.4	30.6	3.8	0.6	0.0	1.3	1.7	0.5	0.6
I-T-W	0.1	2.5	0.4	17.7	66.8	3.0	0.0	0.1	3.3	4.5	1.0	0.8
I-T-L	0.5	1.2	0.0	3.7	4.5	73.9	3.9	0.2	8.4	1.2	1.3	1.2
L-S	0.2	0.4	0.0	0.0	0.0	0.1	77.6	0.4	12.3	0.4	8.0	0.5
L-T-D	6.5	1.0	0.4	0.0	0.1	0.1	0.7	51.8	28.3	1.4	6.3	3.5
L-T-W	0.4	2.2	0.6	0.0	0.1	0.1	3.9	0.5	83.4	0.5	5.8	2.5
L-T-I	0.6	1.1	0.5	0.7	4.1	0.4	5.1	1.2	23.8	47.4	11.0	4.1
L-T-O	0.4	0.4	1.1	0.0	0.0	0.1	13.3	0.6	29.0	0.6	48.6	5.8
L-T-E	0.4	0.6	0.5	0.0	0.1	0.0	0.9	0.3	14.8	0.8	12.8	68.8

Table 3.6. The multiclass confusion matrix for MobileNetV2 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	78.3	0.6	0.4	0.0	0.1	0.0	0.9	5.0	9.9	0.2	2.9	1.8
W-S	0.2	73.4	10.1	0.0	0.3	0.3	0.1	0.4	13.4	0.2	1.2	0.5
W-T-O	0.6	24.2	41.5	0.0	0.4	0.0	0.0	0.6	18.0	0.6	9.5	4.6
I-S	0.0	0.4	0.1	64.7	28.9	2.6	0.4	0.0	0.9	1.3	0.1	0.6
I-T-W	0.0	2.0	0.9	18.5	66.2	3.3	0.0	0.1	2.7	4.9	0.5	1.0
I-T-L	0.2	0.7	0.0	2.2	7.9	73.6	3.4	0.2	7.9	1.7	1.0	1.3
L-S	0.1	0.3	0.0	0.0	0.0	0.1	79.2	0.1	11.9	0.2	7.7	0.4
L-T-D	5.5	0.7	0.4	0.0	0.0	0.1	1.0	53.8	28.2	1.1	5.5	3.7
L-T-W	0.3	1.9	0.5	0.0	0.0	0.1	3.9	0.4	86.5	0.3	4.1	2.0
L-T-I	0.2	1.2	0.1	0.1	3.9	0.4	4.1	1.0	26.1	48.4	9.5	4.9
L-T-O	0.3	0.4	1.2	0.0	0.0	0.1	13.9	0.6	29.5	0.4	47.9	5.6
L-T-E	0.4	0.5	0.6	0.0	0.1	0.1	0.9	0.3	15.3	0.6	10.7	70.7

Table 3.7. The multiclass confusion matrix for VGG16 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	75.5	0.6	0.6	0.0	0.0	0.2	1.1	5.3	9.9	0.0	3.5	3.3
W-S	0.2	66.1	8.0	0.0	0.1	0.2	0.2	0.1	22.3	0.1	1.7	1.0
W-T-O	0.3	23.6	34.2	0.1	0.1	0.2	0.1	0.2	23.5	0.8	10.1	6.7
I-S	0.0	0.4	0.6	59.2	30.1	3.3	1.5	0.0	1.7	0.9	0.5	1.8
I-T-W	0.1	1.8	1.0	15.1	66.5	1.9	0.0	0.1	5.0	5.4	1.4	1.8
I-T-L	0.5	0.0	0.2	4.5	8.1	64.8	3.9	0.0	13.1	1.9	1.5	1.5
L-S	0.1	0.3	0.0	0.0	0.0	0.0	77.0	0.1	14.8	0.3	7.1	0.2
L-T-D	7.3	0.8	0.7	0.0	0.2	0.0	1.0	43.8	33.3	0.6	7.8	4.5
L-T-W	0.3	1.8	0.4	0.0	0.0	0.1	3.8	0.4	86.9	0.2	4.5	1.7
L-T-I	0.7	1.0	0.3	0.7	5.2	0.6	5.6	1.0	29.9	37.2	12.4	5.4
L-T-O	0.3	0.3	0.9	0.1	0.0	0.0	14.7	0.4	35.2	0.5	41.7	5.8
L-T-E	0.6	0.6	0.5	0.0	0.0	0.0	1.0	0.3	19.6	0.3	10.0	67.0

Table 3.8. The multiclass confusion matrix for VGG19 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	73.9	0.5	0.4	0.0	0.1	0.1	1.2	4.4	12.0	0.2	3.6	3.7
W-S	0.4	65.0	7.3	0.0	0.3	0.2	1.0	0.3	21.9	0.3	2.0	1.1
W-T-O	1.1	24.5	31.0	0.1	0.3	0.7	0.1	1.4	24.2	0.8	8.7	7.3
I-S	0.0	0.5	0.0	73.0	19.7	3.1	0.5	0.0	1.4	0.8	0.5	0.5
I-T-W	0.1	1.1	1.1	20.7	61.7	2.4	0.1	0.1	5.1	5.3	1.1	1.2
I-T-L	0.3	1.2	0.2	5.2	5.7	66.3	4.5	0.0	12.0	1.9	1.2	1.5
L-S	0.0	0.5	0.0	0.0	0.0	0.1	76.8	0.0	13.5	0.4	8.5	0.2
L-T-D	6.4	0.5	0.8	0.0	0.1	0.0	1.6	41.2	35.9	0.4	8.2	4.9
L-T-W	0.4	1.8	0.4	0.0	0.0	0.1	4.0	0.4	86.1	0.2	4.7	1.8
L-T-I	0.4	1.6	0.2	0.5	5.1	0.5	5.9	0.3	28.6	37.6	13.5	5.8
L-T-O	0.5	0.3	0.9	0.0	0.0	0.1	15.6	0.5	34.5	0.3	41.5	5.7
L-T-E	0.4	0.7	0.4	0.0	0.0	0.1	1.0	0.3	21.2	0.5	10.9	64.5

Table 3.9. The multiclass confusion matrix for Xception showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	78.4	0.8	0.4	0.0	0.1	0.2	0.9	3.1	10.2	0.2	3.1	2.7
W-S	0.3	68.1	10.7	0.0	0.1	0.0	0.0	0.2	17.7	0.3	1.2	1.2
W-T-O	0.3	23.1	39.0	0.1	0.5	0.2	0.1	0.6	18.6	0.7	7.6	9.3
I-S	0.1	0.9	0.3	57.9	31.6	3.7	0.5	0.0	1.9	1.0	1.3	0.8
I-T-W	0.1	2.2	1.0	14.9	67.7	2.1	0.0	0.5	4.2	5.3	1.2	0.8
I-T-L	0.2	0.5	1.2	2.7	5.6	73.4	3.2	0.3	8.6	1.7	1.9	0.8
L-S	0.1	0.2	0.0	0.0	0.0	0.2	71.7	0.1	13.9	0.4	12.7	0.6
L-T-D	6.7	0.9	0.6	0.0	0.1	0.0	0.7	50.3	28.1	2.8	6.7	3.0
L-T-W	0.4	1.8	0.5	0.0	0.0	0.1	3.5	0.5	83.1	0.5	6.9	2.6
L-T-I	0.2	1.8	0.4	0.4	3.7	0.5	4.0	1.6	22.9	46.2	13.2	5.0
L-T-O	0.6	0.3	1.4	0.0	0.0	0.1	12.8	0.5	28.5	0.7	48.7	6.4
L-T-E	0.3	0.6	0.5	0.0	0.1	0.1	0.8	0.4	14.7	0.6	11.8	70.1

Table 3.10. The multiclass confusion matrix for ResNet50 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	70.8	0.4	0.3	0.0	0.0	0.2	1.2	13.1	9.1	0.2	2.9	1.8
W-S	0.0	63.4	9.1	0.0	0.2	0.2	0.0	8.6	15.7	0.2	1.8	0.8
W-T-O	0.4	22.3	38.5	0.0	0.7	0.1	0.3	3.4	18.0	0.4	8.2	7.7
I-S	0.1	1.2	0.3	67.9	23.8	2.3	0.8	0.0	1.8	0.6	0.9	0.4
I-T-W	0.0	1.6	0.5	17.7	65.7	2.5	0.0	1.2	2.6	5.9	0.9	1.3
I-T-L	0.8	1.0	0.3	2.9	4.7	72.6	4.2	1.0	9.4	1.7	0.5	0.8
L-S	0.1	0.2	0.0	0.0	0.0	0.1	77.6	0.6	12.6	0.4	8.1	0.4
L-T-D	5.7	0.3	0.9	0.0	0.0	0.0	0.7	57.5	24.2	1.7	5.1	3.9
L-T-W	0.4	1.6	0.5	0.0	0.0	0.1	3.1	5.2	80.4	0.5	5.7	2.5
L-T-I	0.6	2.0	0.1	0.3	4.1	0.3	4.3	3.6	20.7	48.7	11.0	4.3
L-T-O	0.2	0.4	0.9	0.0	0.0	0.0	13.2	1.7	28.5	0.8	48.2	5.9
L-T-E	0.6	0.7	0.5	0.0	0.0	0.1	0.8	1.4	14.8	0.8	10.1	70.3

Table 3.11. The multiclass confusion matrix for ResNet101 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	77.7	0.5	0.2	0.0	0.1	0.3	1.1	4.5	9.7	0.1	3.5	2.3
W-S	0.4	66.6	10.1	0.0	0.2	0.2	0.1	0.3	18.3	0.1	2.2	1.5
W-T-O	0.3	24.5	41.8	0.0	0.4	0.1	0.0	0.4	15.6	0.6	7.5	9.0
I-S	0.1	0.5	0.1	63.1	28.7	2.2	0.5	0.0	1.5	1.3	1.3	0.6
I-T-W	0.1	1.8	0.8	20.1	63.2	3.5	0.0	0.2	2.4	5.1	0.6	2.2
I-T-L	0.5	1.0	0.2	3.5	3.9	58.2	2.7	0.2	10.6	1.9	2.0	15.3
L-S	0.1	0.4	0.0	0.0	0.0	0.1	75.0	0.1	12.8	0.6	8.9	2.0
L-T-D	7.3	0.7	0.6	0.0	0.1	0.0	1.0	53.1	26.2	1.8	5.8	3.5
L-T-W	0.3	1.7	0.5	0.0	0.0	0.1	3.5	0.6	81.5	0.4	6.5	4.9
L-T-I	0.2	1.8	0.2	0.4	3.4	0.2	4.1	1.6	21.9	47.9	12.1	6.2
L-T-O	0.5	0.4	1.0	0.0	0.0	0.1	12.5	0.6	27.8	0.7	48.6	8.0
L-T-E	0.8	0.6	0.6	0.0	0.1	0.1	0.7	0.3	13.6	0.6	11.1	71.5

Table 3.12. The multiclass confusion matrix for ResNet152 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	77.1	0.6	0.6	0.0	0.1	0.1	1.0	4.2	10.4	0.2	3.4	2.3
W-S	0.6	60.5	10.9	0.0	0.1	0.3	0.0	0.7	25.0	0.3	1.1	0.6
W-T-O	0.4	20.3	42.3	0.0	0.2	0.0	0.0	0.4	23.7	0.6	8.2	4.1
I-S	0.1	0.9	0.1	62.5	27.9	3.3	1.0	0.0	1.8	1.7	0.5	0.1
I-T-W	0.1	1.8	0.8	18.6	64.8	2.6	0.3	0.5	2.9	5.4	1.2	1.2
I-T-L	0.2	1.0	0.3	2.5	5.2	72.2	3.9	0.2	9.9	1.9	1.7	1.0
L-S	0.1	0.3	0.0	0.0	0.0	0.1	78.5	0.3	12.7	0.3	7.4	0.4
L-T-D	5.8	0.5	0.4	0.0	0.2	0.1	0.9	54.5	26.2	1.8	6.4	3.2
L-T-W	0.3	1.5	0.6	0.0	0.0	0.1	3.8	0.5	85.4	0.5	5.0	2.3
L-T-I	0.1	0.9	0.1	0.4	3.8	0.7	4.7	1.5	23.5	46.7	12.6	5.1
L-T-O	0.4	0.4	0.9	0.0	0.1	0.0	14.0	0.8	29.9	0.5	46.9	6.0
L-T-E	0.4	0.5	0.4	0.0	0.0	0.1	0.7	0.7	15.6	0.7	9.9	70.9

Table 3.13. The multiclass confusion matrix for DenseNet121 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	79.2	0.5	0.1	0.0	0.2	0.3	1.1	4.1	9.0	0.3	3.1	2.2
W-S	0.2	66.4	8.8	0.0	0.3	0.1	0.0	0.4	21.5	0.3	1.4	0.6
W-T-O	0.5	21.1	35.2	0.0	0.4	0.0	1.0	0.2	24.1	0.8	11.1	5.7
I-S	0.1	0.3	0.6	60.8	29.2	2.3	0.5	0.0	2.3	0.9	1.8	1.2
I-T-W	0.2	1.8	0.4	20.3	64.9	3.1	0.0	0.1	2.8	4.6	0.8	1.1
I-T-L	0.2	0.7	0.3	3.0	4.9	74.6	3.2	0.3	8.6	1.9	1.3	1.0
L-S	0.1	0.1	0.0	0.0	0.0	0.2	78.3	0.3	11.9	0.2	8.3	0.7
L-T-D	6.0	0.7	0.4	0.0	0.2	0.1	0.8	53.0	27.2	2.4	5.7	3.4
L-T-W	0.4	1.6	0.5	0.0	0.0	0.1	3.9	0.5	84.4	0.5	5.8	2.2
L-T-I	0.1	1.2	0.3	0.7	4.1	0.7	4.6	1.1	23.6	48.8	10.7	4.1
L-T-O	0.5	0.4	0.8	0.0	0.0	0.1	13.8	0.6	29.3	0.6	48.6	5.2
L-T-E	0.7	0.6	0.4	0.0	0.1	0.2	0.7	0.4	15.8	0.8	10.7	69.6

Table 3.14. The multiclass confusion matrix for DenseNet169 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	75.0	0.5	0.1	0.0	0.1	0.1	0.8	4.0	13.7	0.1	3.3	2.2
W-S	0.3	68.9	8.8	0.0	0.1	0.2	0.0	0.2	18.5	0.2	1.8	1.1
W-T-O	0.3	24.1	36.9	0.0	0.2	0.1	0.6	0.4	19.6	0.4	7.1	10.4
I-S	0.0	0.8	0.1	58.5	32.8	2.4	0.4	0.1	2.3	1.4	1.0	0.1
I-T-W	0.2	2.0	0.9	17.8	64.1	3.9	0.0	0.1	5.5	3.9	1.0	0.8
I-T-L	0.3	0.5	0.0	2.9	5.9	72.9	3.0	0.2	9.8	1.9	1.9	0.8
L-S	0.2	0.3	0.0	0.0	0.0	0.2	78.1	0.2	12.6	0.3	7.6	0.5
L-T-D	7.0	0.9	0.6	0.0	0.2	0.1	0.8	50.8	28.5	2.0	5.7	3.5
L-T-W	0.4	1.8	0.5	0.0	0.1	0.1	3.6	0.6	84.3	0.4	5.4	2.9
L-T-I	0.2	1.7	0.1	0.3	4.8	0.2	3.6	1.8	24.3	45.9	12.6	4.4
L-T-O	0.5	0.3	0.9	0.0	0.0	0.1	13.6	0.7	30.1	0.5	46.9	6.3
L-T-E	0.6	0.6	0.5	0.0	0.0	0.1	0.7	0.5	18.0	0.6	10.2	68.2

Table 3.15. The multiclass confusion matrix for DenseNet201 showing the image classification accuracies (%) during inference on the ExoNet database. The columns and rows are the predicted and labelled classes, respectively.

	D-T-L	W-S	W-T-O	I-S	I-T-W	I-T-L	L-S	L-T-D	L-T-W	L-T-I	L-T-O	L-T-E
D-T-L	75.6	0.9	0.2	0.0	0.0	0.4	0.7	5.1	10.0	0.1	4.0	2.8
W-S	0.3	67.9	12.1	0.0	0.2	0.1	0.0	0.3	16.2	0.6	1.7	0.5
W-T-O	2.4	23.6	37.8	0.0	0.4	0.2	0.0	1.5	19.9	0.4	8.3	5.5
I-S	0.1	0.0	0.3	59.4	33.7	2.7	0.1	0.1	1.5	1.2	0.9	0.0
I-T-W	0.2	1.9	1.0	14.2	68.9	3.0	0.1	0.1	2.8	5.6	1.2	1.0
I-T-L	0.2	0.5	0.2	4.0	5.7	72.2	2.2	0.2	10.3	2.0	1.3	1.2
L-S	0.0	0.2	0.0	0.0	0.0	0.1	76.0	0.2	13.4	0.6	8.9	0.6
L-T-D	6.4	0.7	0.8	0.0	0.2	0.1	1.2	51.4	27.0	2.7	5.0	4.6
L-T-W	0.4	1.8	0.9	0.0	0.0	0.1	3.5	0.6	82.3	0.5	7.0	2.8
L-T-I	0.7	1.5	0.4	0.5	4.5	0.4	5.1	1.3	23.9	47.2	10.8	3.7
L-T-O	0.5	0.4	1.4	0.0	0.0	0.1	12.6	0.8	29.6	0.8	47.4	6.4
L-T-E	0.8	0.6	0.4	0.0	0.1	0.1	0.9	0.6	13.5	0.7	12.0	70.4

Tables 3.3-3.15 show the resulting multiclass confusion matrix for each CNN architecture. The matrix columns and rows are the predicted and labelled classes, respectively. The diagonal elements are the classification accuracies (%) for each environment class during inference and the nondiagonal elements are the misclassification percentages; the darker shades represent higher classification accuracies. Despite slight numerical differences, most of the CNN architectures displayed a similar interclass trend. The networks most accurately predicted the "level-ground transition wall/door" (L-T-W) class, with an average accuracy of 84.3% ± 2.1%, followed by the "level-ground steady" (L-S) class with an average accuracy of 77.3% ± 2.1% and the "decline stairs transition level-ground" (D-T-L) class with an average accuracy of 76.6% ± 2.4%. These results could be attributed to the class imbalances in the dataset such that there were significantly more images of L-T-W environments compared to other classes. However, some classes with limited images showed relatively good classification performance. For example, the "incline stairs transition level-ground" (I-T-L) class contains only 1.2% of the ExoNet database but achieved 71.2% ± 5.1% average classification accuracy. Not surprisingly, the least accurate predictions were for the two environment classes that contain "other" features – i.e., the "wall/door transition other" (W-T-O) class with an average accuracy of 38.3% ± 3.5% and the "level-ground transition other" (L-T-O) class with an average accuracy of 46.8% ± 2.4%. These lower prediction accuracies are likely due to the increased noise and randomness of the environments and/or objects within the images.

The development of these deep convolutional neural networks has traditionally focused on improving classification accuracy, often leading to more accurate yet inefficient algorithms with greater computational and memory storage requirements [133]. These design features can be especially problematic for deployment on mobile and embedded systems, which inherently have limited operating resources. Despite advances in computing devices like graphics processing units (GPUs), the current embedded systems in robotic leg prostheses and exoskeletons would struggle to support the architectural and computational complexities typically associated with deep learning for computer vision. For onboard real-time inference, the ideal convolutional neural network would achieve high image classification accuracy with minimal parameters, computing operations, and inference time. Motivated by these principles, the operational performances of the benchmarked CNN architectures $(\mathcal{N})$ on the ExoNet database were quantitatively evaluated and compared using a balanced metric called "NetScore" [134]:

$$\Omega(\mathcal{N}) = 20log\left(\frac{a(\mathcal{N})^{\alpha}}{p(\mathcal{N})^{\beta}m(\mathcal{N})^{\gamma}}\right) \quad (2)$$

where $a(\mathcal{N})$ is the image classification accuracy during inference (0-100%); $p(\mathcal{N})$ is the number of parameters in millions; $m(\mathcal{N})$ is the number of multiply–accumulates in billions; and α, β, and γ are coefficients that control the effects of the image classification accuracy, and the architectural and computational complexities on the NetScore (Ω), respectively. The coefficients wet set to

$\{\alpha = 2;\ \beta = 0.5;\ \gamma = 0.5\}$ to better emphasize the classification accuracy while partially considering the parameters and computing operations since CNNs with low accuracy are less practical, regardless of the size and speed. Note that the NetScore does not explicitly account for inference time. The number of parameters and multiply–accumulates are assumed to be representative of the architectural and computational complexities, respectively, both of which are inversely proportional to the NetScore (Table 3.2). Although the ResNet152 network achieved one of the highest classification accuracies on the ExoNet database (71.6% accuracy), it received the lowest NetScore $(\Omega = 46)$ due to the disproportionally large number of parameters (i.e., containing more parameters than any other CNN architecture). Interestingly, the EfficientNetB0 network did not receive the highest NetScore $(\Omega = 72.6)$ despite achieving the highest image classification accuracy (73.2% accuracy) and the architecture having been optimized using a neural architecture search to maximize the classification accuracy while minimizing the number of computing operations [122].

The MobileNetV2 network developed by Google [125], which uses depthwise separable convolutions, received the highest NetScore $(\Omega = 76.2)$, therein demonstrating the best balance between the classification accuracy (72.9% accuracy) and the architectural and computational complexities. Depthwise separable convolutions factorize standard convolutions into depthwise convolutions and 1×1 convolutions called "pointwise" convolutions. The depthwise convolution layer spatially convolves a filter over each input channel separately. The pointwise convolution layer then convolves a 1×1 filter across each input channel to combine the outputs from the depthwise convolution layer and generate a new feature representation. By using depthwise separable convolutions (i.e., splitting the filtering and combining processes into separate layers), MobileNets tend to have significantly fewer parameters and computing operations compared to other state-of-the-art deep convolutional neural networks used in computer vision (Table 3.2).

The researchers at Google previously demonstrated the ability of MobileNetV2 to perform onboard real-time inference on a mobile computing device (i.e., ~75 ms per image on a CPU-powered Google Pixel 1 smartphone) [125]. However, the classification system developed here could theoretically generate even faster runtimes since 1) the smartphone camera that was used (i.e., the iPhone XS Max) has an onboard GPU; and 2) the size of the final densely connected layer of the MobileNetV2 architecture was reduced from 1,000 outputs, as originally used for ImageNet, to 12 outputs, for the ExoNet database. Compared to traditional CPUs, GPUs have many more core processors, which can permit faster and more efficient CNN computations through parallel computing [55]. Moving forward, the existing CPU embedded systems in robotic leg prostheses and exoskeletons should be used for the locomotion mode recognition based on neuromuscular-mechanical data, which is less computationally expensive, and a supplementary GPU computing device for the vision-based environment classification; these recommendations concur with those recently proposed by Huang and colleagues [36].

It is important to emphasize that the use of state-of-the-art deep convolutional neural networks was made possible because of the ExoNet database. In addition to being open-source, the large scale and diversity of ExoNet significantly distinguishes itself from all previous datasets (see Table 2.1). ExoNet contains ~923,000 labelled images. In comparison, the previous largest dataset, developed by Varol's research lab [44], contained ~402,000 images. Whereas previous datasets have included fewer than six environment classes, the most common being level-ground terrain and incline and decline stairs, the ExoNet database includes a novel hierarchical labelling architecture with 12 distinct classes. These differences can have important practical implications since deep learning requires significant and diverse training data to prevent overfitting and promote generalization [55]. Furthermore, the quality of the ExoNet images (1280×720) is considerably higher than previous datasets (e.g., 224x224 and 320x240). Poor image resolution has been shown to decrease the environment classification accuracy [36], [42]. Although higher resolution images can increase the onboard computational and memory storage requirements, using efficient CNN architectures with fewer computing operations like EfficientNets [122] can allow for processing larger images for relatively similar computational cost.

As robotic leg prostheses and exoskeletons begin to transition out of research laboratories and into real-world environments, large-scale and challenging datasets like ExoNet are needed to support the development of next-generation image classification algorithms for environment-adaptive locomotor control. In addition to robotic leg prostheses and exoskeletons, applications of ExoNet could extend to humanoids, autonomous legged robots, powered wheelchairs, assistive devices for persons with visual impairments, in addition to fundamental science research on legged locomotion in real-world environments.

Building on the results of this chapter, the ExoNet database was used to inform the energy regeneration research in Chapter 4. Previous studies of robotic leg prostheses and exoskeletons with regenerative actuators have focused on steady-state level-ground walking [32]. However, as illustrated in Figure 3.6, steady-state level-ground terrain, labelled as L-S, represents a relatively small percentage of real-world walking environments (8.4%). This observation concurs with [112], who showed that, in real-world community mobility, steady-state locomotion is generally short lived and separated by frequent transitions between different states (e.g., ~40% of walking bouts are less than 12 consecutive steps). There are other locomotor activities of daily living like stair descent and stand-to-sit movements that involve net negative mechanical work and thus present an opportunity for energy regeneration. Decline stairs, labelled as D-T-L in the ExoNet database, represent 3.1% of real-world walking environments whereas seat environments, labelled as L-T-E, appear much more frequently (13%). These differences in the prevalence of stairs versus seat environments may be even larger in the residential communities of persons with mobility impairments (e.g., long-term care facilities). Accordingly, Chapter 4 explored the potential for energy regeneration during stand-to-sit movements.

3.3 Chapter Summary

In this chapter, a multi-generation environment classification system powered by computer vision and deep learning was developed for robotic leg prostheses and exoskeletons. Taking inspiration from the human vision-locomotor control system, environment sensing and classification could improve the automated high-level control and decision-making of these wearable robotic devices by predicting the oncoming walking environment prior to physical interactions, therein allowing for more accurate and robust locomotion mode transitions (e.g., switching between level-ground walking and stair ascent). As a proof-of-concept, Section 3.1 presented the development and evaluation of a preliminary environment recognition system. Images of indoor and outdoor real-world walking environments were collected around the University of Waterloo campus using a wearable camera and a convolutional neural network was designed and trained to recognize level-ground terrain, and incline and decline stairs, therein achieving 94.85% image classification accuracy. This preliminary environment classification system demonstrated the feasibility of using deep learning for this novel computer vision application.

Building on the preliminary system design, Section 3.2 outlined the development and evaluation of a second-generation environment recognition system. Over 5.6 million images of indoor and outdoor real-world walking environments were collected using a lightweight wearable camera, of which ~923,000 images were annotated using a hierarchical labelling architecture with 12 individual classes, known as the ExoNet database. Available publicly through the IEEE DataPort repository, ExoNet offers an unprecedented communal platform to train, develop, and compare next-generation image classification algorithms for visual recognition of legged locomotion environments.

Over a dozen state-of-the-art deep convolutional neural networks were then trained and tested on ExoNet for image classification and automatic feature engineering, including: EfficientNetB0; InceptionV3; MobileNet; MobileNetV2; VGG16; VGG19; Xception; ResNet50; ResNet101; ResNet152; DenseNet121; DenseNet169; and DenseNet201. The benchmarked CNN architectures and their environment classification predictions were quantitatively evaluated and compared using an operational metric called NetScore, which balances the classification accuracy with the architectural and computational complexities (i.e., important for onboard real-time inference with mobile computing devices). The comparative analyses showed that the EfficientNetB0 network achieved the highest test accuracy; VGG16 the fastest inference time; and MobileNetV2 the best NetScore and least number of parameters and computing operations, which can inform the optimal architecture design or selection depending on the desired performance. This second-generation environment sensing and classification system, part of which was recently published in the 2021 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) [135], [136], provides a large-scale benchmark and reference for future research in environment-adaptive locomotor control.

4. ENERGY REGENERATION

To reiterate, the overall objective of this thesis research is to support the transition of robotic leg prostheses and exoskeletons from research laboratories to real-world environments via developments in environment-adaptive control (Chapter 3) and energy-efficient actuation. In this chapter, an energy regeneration system was developed and evaluated using mathematical and computational models of human and wearable robotic systems to simulate energy regeneration and storage during locomotor activities of daily living.

4.1 Biomechanics for Regeneration

4.1.1 Introduction

As reviewed in Chapter 2, backdriveable actuators with energy regeneration can improve the energy efficiency and extend the battery-powered operating durations of robotic leg prostheses and exoskeletons. Previous related studies have focused on steady-state level-ground walking, specifically regenerating energy during late swing knee extension [69], [70], [86]–[89], [91], [94], [97]–[100], [100]–[102], [104], [106]–[111]. In real-world community mobility, however, steady-state locomotion is generally short lived and separated by frequent transitions between different states [112]. This trend concurs with the research in Chapter 3, which showed that a relatively small percentage of real-world legged locomotion environments consist of steady-state level-ground terrain (see Figure 3.6). Moreover, given that targeted users of these wearable robotic devices (i.e., older adults and/or individuals with physical disabilities) tend to walk slower and take fewer steps per day [2], and that energy regeneration and efficiency both positively relate to walking speed [83], [88]–[90], [104], [106], [107], [113], the potential for energy regeneration from only steady-state level-ground walking is relatively limited.

Real-world community mobility involves many stops and starts, often beginning and ending with a seated or standing posture [112]. Sit-to-stand and stand-to-sit movements are common daily locomotor activities of persons with mobility impairments and are prerequisites to walking. Although several robotic leg prostheses and exoskeletons have been designed and evaluated for sitting and standing transitions [137]–[147], these devices did not include regenerative actuators. Regenerating energy while sitting down is a relatively unexplored and potentially viable method to help recharge the onboard batteries by recovering some of the otherwise dissipated energy during negative mechanical work via backdriving the electromagnetic motor.

As a simple approximation, the biomechanical energy theoretically available for electrical energy regeneration while sitting down can first be estimated as the change in mechanical energy, specifically gravitational potential energy, of the center of mass of the total body system modelled as a point mass. The change in total body mechanical energy during stand-to-sit movements is

representative of the mechanical work (W) done on the system. This work-energy principle can be described by

$$W = \Delta E_{gpe} \tag{3}$$

$$E_{gpe} = Mgy_0 \tag{4}$$

where E_{gpe} is the gravitational potential energy of the total body center of mass, M is the body mass, g is the gravitational acceleration (9.81 m/s^2), and y_0 is the height of the total body center of mass above the horizontal reference axis. Since the shank and feet are assumed to be relatively stationary during stand-to-sit movements, these preliminary calculations model the human as an inverted double pendulum with only HAT and thigh segments (i.e., M equals the combined mass of the HAT and bilateral thigh segments). Anthropometric proportions from cadaver research [59] were used to estimate the body segment parameters from total body mass (M_{total}) and height, which were assumed to be 78 kg and 1.8 m, respectively. The mass and length of each segment was estimated as a proportion of the total body mass and height, respectively, and the center of mass position was estimated as a proportion of the segment length. The mass (m_s) of each segment was calculated by

$$m_s = P_s M_{total} \tag{5}$$

where P_s is the segment's mass proportion, which are 0.678 and 0.200 for the HAT and bilateral thigh segments, respectively [59]. A similar calculation was used to estimate the segment lengths. The length proportions for the HAT and thigh segments are 0.470 and 0.245, respectively [59]. The center of mass position of each segment was estimated as a proportion of the distance from the proximal to the distal endpoints $\left(R_{proximal}\right)$. These equations were used to calculate the vertical coordinates (i.e., height) of the center of mass position (y_{cm}) of each segment in both standing and seated postures:

$$y_{cm} = y_{proximal} + R_{proximal}\left(y_{distal} + y_{proximal}\right) \tag{6}$$

where $y_{proximal}$ and y_{distal} are the coordinates of the proximal and distal endpoints, respectively, which were estimated based on the segment lengths of the human system when standing and seated (i.e., the thigh and horizontal reference axis are parallel). The proportions for the center of mass position of the HAT and thigh segments are 1.142 and 0.433, respectively [59]. The height of the center of mass of the total body system in standing and seated postures was calculated by

$$y_0 = \frac{(m_s^{HAT} y_{cm}^{HAT}) + (m_s^{THIGH} y_{cm}^{THIGH})}{M} \tag{7}$$

where M equals 68.5 kg. The difference in y_0 between standing and seated postures was 0.398 m (i.e., the distance travelled by the total body center of mass). Based on Equation (4), the estimated change in mechanical energy while sitting down is ~267 J, which may be representative of the

biomechanical energy available for electrical energy regeneration. Although these equations for estimating mechanical energy are relatively simple since only linear kinematics and body segment parameters are needed, this method has several limitations [59]. A more informative and accurate method uses inverse dynamics to compute the net moments of force at individual joints, which can be used to calculate the joint mechanical work and power in order to study the distribution of energy generation and absorption throughout the lower-limbs. Accordingly, this research study quantified the lower-limb joint mechanical work and power during stand-to-sit movements using parameter identification and inverse dynamic simulations of subject-specific optimized biomechanical models to estimate the mechanical energy theoretically available for electrical energy regeneration.

4.1.2 Motion Capture Experiments

Nine subjects were recruited and provided informed written consent (height: 180 ± 4 cm; total body mass: 78 ± 7 kg; age: 25 ± 3 years; sex: male). Each participant performed 20 sit-to-stand and stand-to-sit movements at self-selected rates while lower-limb kinematics and ground reaction forces were experimentally measured using motion capture cameras and force plates, respectively (Figure 4.1). Separate force plates were used to measure the ground reaction forces underneath the chair and feet. The seat height was ~46 cm. The motion capture cameras (Optotrak) provided 3D measurements of active marker positions in the global coordinates. Active marker systems are generally considered the gold standard in human movement biomechanics. The motion capture cameras and force plates were sampled at 100 Hz and 300 Hz, respectively. For tracking individual body segment positions in the sagittal plane, virtual markers were digitized overlying palpable anatomical landmarks on the right lower-limb, including the lateral malleolus, lateral femoral and tibial condyles, and greater trochanter. These virtual marker positions correspond with those recommended by the International Society of Biomechanics [148]. This study was approved the University of Waterloo Office of Research Ethics.

Missing marker data were estimated using cubic spline interpolations. The ankle and hip joint centers were assumed at the lateral malleolus and greater trochanter marker positions, respectively. The estimated knee joint center was the midpoint between the lateral femoral and tibial condyle markers. Piecewise cubic Hermite interpolating polynomials were used to resample and time-normalize (0-100%) the kinematic measurements. Average line vectors between the ankle and knee joint centers, and the knee and hip joint centers, defined the shank and thigh body segment lengths, respectively. Inverse kinematics was used to convert the virtual marker positions to independent joint coordinates through vector algebra. The ankle joint angle was the relative angle between the shank and horizontal axis. The relative angle between the shank and thigh segments defined the knee joint angle. Given the relative rotations between the pelvis and head-arms-trunk (HAT) segment (i.e., pelvis tilting), the experimentally measured pelvis marker-cluster rotations differed from HAT segment rotations. Therefore, the HAT segment was assumed vertical

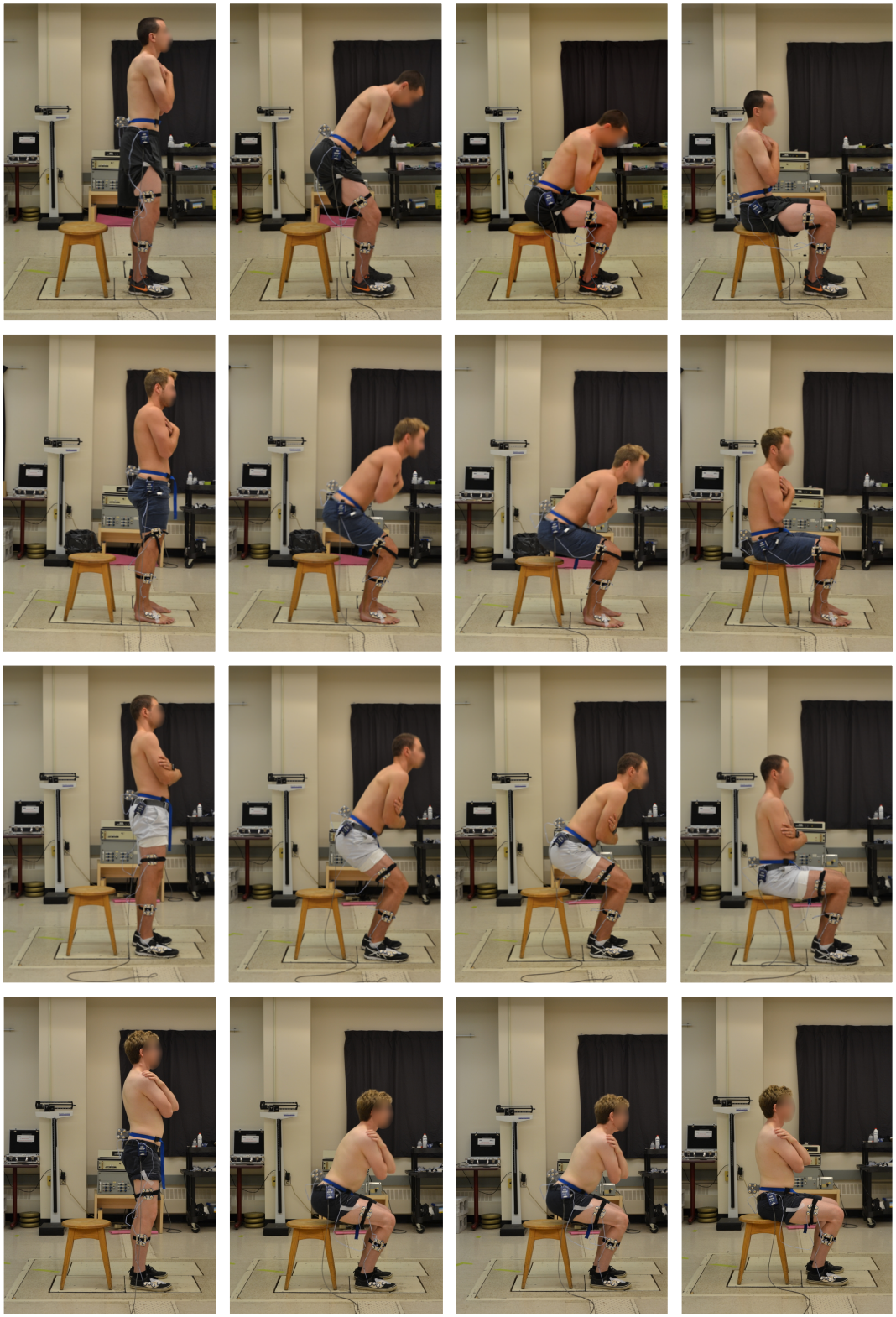

Figure 4.1. Examples of the stand-to-sit movements used for energy regeneration, the biomechanics of which were experimentally measured using motion capture cameras and force plates. Nine healthy young adults (n=9) were tested in total.

when standing (i.e., initial posture) and seated (i.e., final posture) and the pelvis angle rotations were assumed to progress linearly throughout the movement. The relative angle between the thigh and HAT segments defined the hip angle.

Joint angles were filtered using a 10th-order low-pass Butterworth filter with a 5 Hz cut-off frequency and zero-phase digital filtering [59]. The joint rotational speeds and accelerations were calculated by numerically differentiating the joint angles. Similar to Gao et al. [137], the pelvis translational velocities, which were estimated from the low-back marker-cluster shown in Figure 4.1, were used to segment sitting and standing movements. These kinematic measurements were filtered using a 10th-order low-pass Butterworth filter with a 3 Hz cut-off frequency, and zero-phase digital filtering and moving average smoothing filtering. The sitting and standing movements were segmented when the pelvis translational speeds exceeded a percentage of their maximum values (i.e., the most common being 15% and 20%), which were estimated through trial-and-error simulations. The force plate measurements were filtered using a 10th-order low-pass Butterworth filter with a 30 Hz cut-off frequency, and zero-phase digital filtering [59]. Piecewise cubic Hermite interpolating polynomials were used to time-normalize the force plate measurements.

4.1.3 Biomechanical Model

The human biomechanical system was modelled using MapleSim software (Maplesoft, Canada), which uses linear graph theory and principles of mechanics to generate a unified representation of the system topology and modelling coordinates. The multibody biomechanical model consisted of a 2D sagittal-plane, inverted triple-pendulum with shank, thigh, and HAT rigid body segments (see Figure 4.2). Since the foot marker position remained relatively unchanged during the motions (i.e., maximum horizontal and vertical displacements of 1.27 cm and 0.7 cm, respectively), the foot segment was modelled as rigidly fixed to the ground frame. The ankle, knee, and hip joints were modelled as idealized revolute kinematic pairs. Biological passive joint torques, including stiffness and damping, were ignored since ideal joints were assumed for modelling an exoskeleton or prosthetic system. Frictionless joints ensured that the joint mechanical work on the system was due to the net joint torques.

The biomechanical model had 3 degrees-of-freedom and mathematically described by 3 generalized coordinates (i.e., the ankle, knee, and hip joint angles) with zero algebraic constraints. Assuming the foot segment was rigidly fixed to the ground and had relatively small mass, the ground reaction forces underneath the foot corresponded to the ankle joint reaction forces, and the ground reaction moments were offset by the ankle position relative to the center of pressure (COP). The center of mass of each body segment was defined relative to the distal joint center, and the moment of inertia was defined about the center of mass position. The measured ground reaction forces underneath the seat were applied to the biomechanical model buttocks when seated. The multibody system equations used for inverse dynamics consisted of only algebraic

equations. MapleSim automatically generated these equations symbolically using graph-theoretic algorithms, therein allowing for multidomain systems modelling (e.g., humans interacting with wearable robotic devices) and computationally efficient dynamic simulations.

4.1.4 Simulation and Parameter Identification

The biomechanical model was driven using the experimental joint kinematics and seat forces. The ankle, knee, and hip joint torques (τ_j), and the ground reaction forces and moment underneath the feet, were calculated from inverse dynamics. In conventional "bottom-up" inverse dynamics, the joint reaction forces and torques are solved segment-by-segment, starting from ground and moving sequentially upward. However, due to limitations in the measured kinematics, system parameters, and/or unmodeled dynamics, the calculated forces and accelerations on the final body segment (i.e., the HAT segment) do not satisfy the system equations [149], [150]. In contrast, the biomechanical model in this study was driven using the experimental joint kinematics in a multibody dynamics simulation and thus could be considered more dynamically consistent.

The simulated ground reaction forces from inverse dynamics were compared with those experimentally measured while the body segment inertial parameters of the biomechanical model were concurrently optimized to minimize the differences. Although subject-specific body segment parameters can be estimated using medical imaging [151], [152] and/or anthropometric proportions from cadaver research, this study used system parameter identification for better dynamical consistency (Figure 4.2). Dynamic parameter identification can be subcategorized into offline methods (e.g., experimental measurements and system input-output matching) and online methods (e.g., nonlinear adaptive control). The offline parameter identification used here involved constrained nonlinear programming (Fmincon, MATLAB) and an interior-point algorithm to estimate the body inertial segment parameters (i.e., mass, center of mass, and moment of inertia) of the HAT, thigh, and shank. The optimization searched for the system parameters that minimized the sum of squared differences in ground reaction forces (GRF) and moments (GRM) between the experimental measurements (m) and inverse dynamic simulations (s) at each time step (i). The optimization multiobjective cost function was:

$$J = \sum_i w_1 (GRF_m[i] - GRF_s[i])^2 \times \frac{1}{(M_{total})^2} + \sum_i w_2 \left(\left(GRM_m[i] + GRM_{offset} \right) - GRM_s[i] \right)^2 \times \frac{1}{(H \times BW)^2} \quad (8)$$

where the GRF vector included both horizontal (GRF_x) and vertical $\left(GRF_y\right)$ components, GRM was about the z-axis, M_{total} is total body mass, coefficient $H = 1$ m, and GRM_{offset} compensated for the horizontal distance between the ankle and foot center of pressure according to $\left\{GRM_{offset} = GRF_y \times COP_x - GRF_x \times COP_y\right\}$ with COP_x and COP_y being the estimated average positions of the foot COP relative to the ankle. The optimization variables were the mass, center

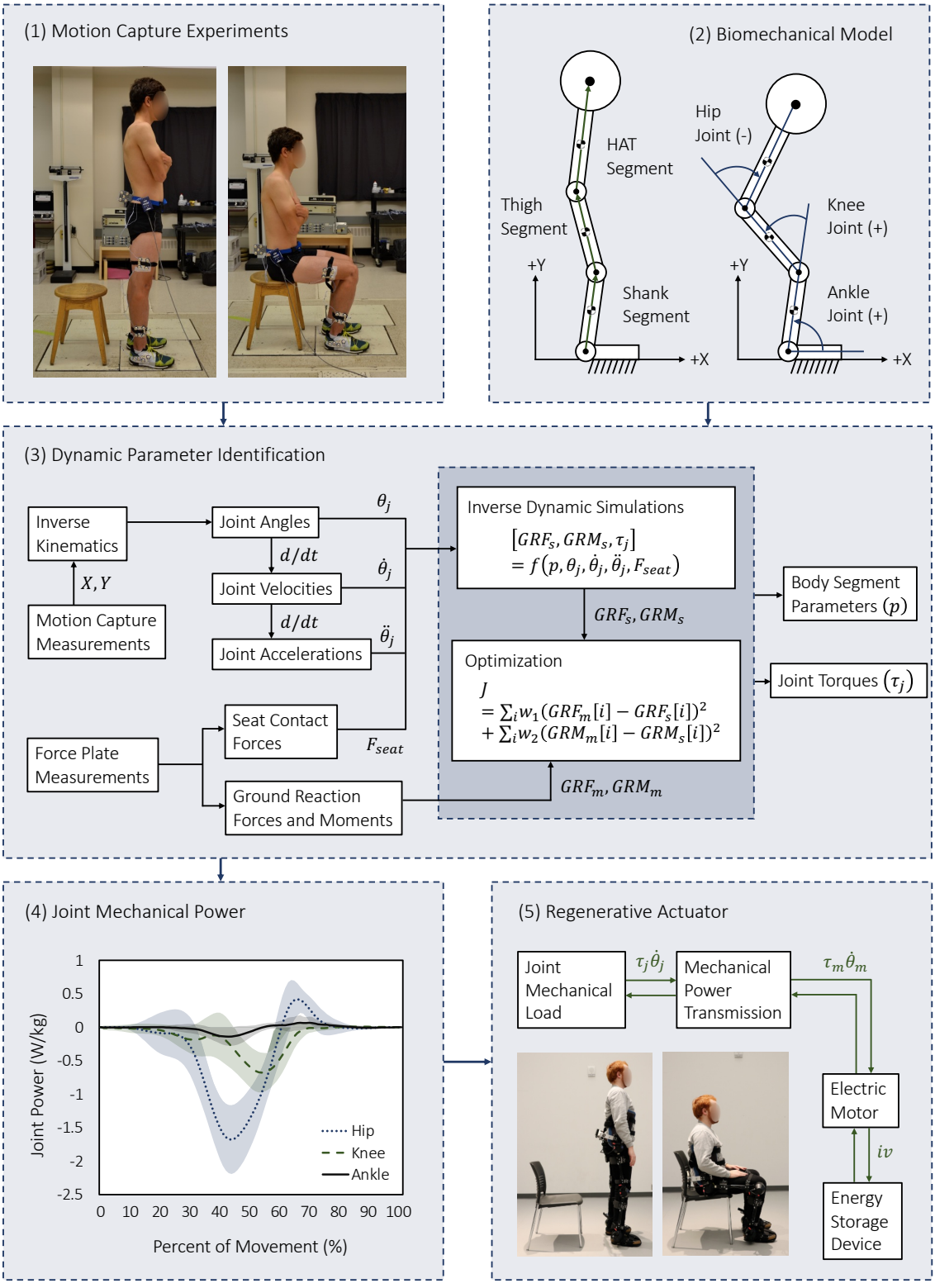

Figure 4.2. Schematic of the experimental and computational methods used to evaluate the stand-to-sit biomechanics for energy regeneration, including (1) the motion capture experiments; (2) biomechanical model; (3) dynamic parameter identification; (4) joint power analyses; and (5) a regenerative actuator model used to convert human mechanical power $(\tau_j \dot{\theta}_j)$ into electrical power (iv) while sitting down.

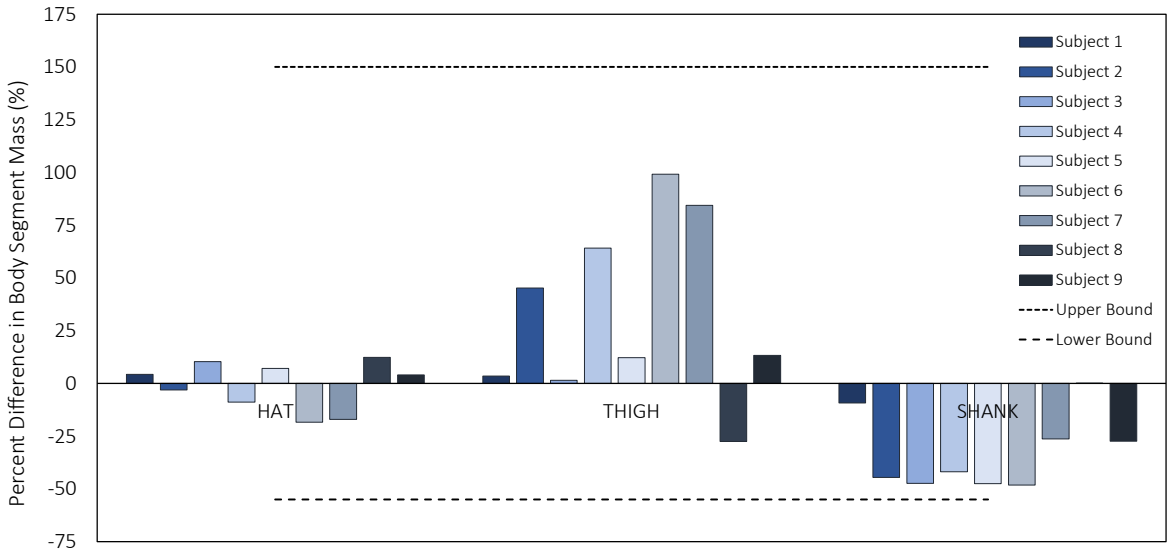

Figure 4.3. The percent difference in HAT, thigh, and shank body segment mass between the system parameter identification (ID) and anthropometric proportions from cadaver research [59] scaled to the measured total body mass of each subject. The positive and negative values represent an overestimation and underestimation in body segment mass, respectively, from the parameter ID compared to the anthropometric proportions. The upper and lower bounds in the optimization are also displayed, the values of which are described in the text.

of mass, and moment of inertia of the shank, thigh, and HAT segments. Other variables included "$seat_{offset}$" and the COP_x. $Seat_{offset}$ was the horizontal distance between the biomechanical model buttocks (i.e., vertical seat force point-of-application) and the hip joint center. Although the COP position moved underneath the base of support, the ground reaction forces were not used to drive the biomechanical model but rather to validate the system parameter identification. The best average position, denoted as parameter COP_x in the optimization algorithm, was estimated by the parameter identification. The COP_y was the ankle marker height. The optimization was constrained by setting 1) lower and upper bounds on individual variables, and 2) the sum of the body segment masses equaled to the measured total body mass. Initial guesses were taken from human anthropometrics and/or were the midpoints between the upper and lower bounds.

The upper and lower bounds for the body segment parameters in the optimization were 3.5 and 0.55 times those estimated from anthropometric proportions from cadaver research [59], respectively, based on the total height and body mass of individual subjects. For example, as illustrated in Figure 4.3, the estimated mass of the HAT, thigh, and shank segments for each subject fall within the upper and lower bounds. Each term in the multiobjective cost function had equal weights. Stopping criteria for the step size and objective function changes were both 1e-14 between iterations. Once the optimal system parameters were found, the joint mechanical powers for the ankle, knee, and hip were calculated from the net joint torques and angular velocities

$\{P_j = \tau_j \dot{\theta}_j\}$ and integrated over time estimate the joint mechanical energy generated and absorbed while sitting down. In other words, the mechanical power provided by the net moment of force, sometimes called the joint moment power, is the product of the net moment of force times the difference between the angular velocities of the two segments that make up the joint. The positive and negative joint mechanical work at each time step were summed to calculate the total positive and negative joint mechanical work, respectively. Using inverse dynamics to calculate the joint mechanical power, and thus mechanical work, can more accurately represent the human musculotendon work compared to traditional methods like external and internal mechanical work calculations [63].

It is important to emphasize that for the system parameter identification, only some of the inertial parameters theoretically affect the system dynamics, while others have an effect in linear combinations. Dynamic parameters can be subclassified into fully-identifiable, identifiable in linear combinations, and unidentifiable. Because the body segment parameters of the biomechanical model are redundant, the inertial parameters of individual body segments (i.e., mass, center of mass, and moment of inertia) cannot be uniquely identified from the experimental measurements alone. Consequently, the individual body segment masses in Figure 4.3 were not explicitly identified. The non-redundant and identifiable inertial parameters, called base inertia parameters, can define the biomechanical model uniquely and sufficiently describe the system dynamics. However, to calculate the base inertia parameters, the multibody system equations would need to be rewritten in linear form, which is beyond the scope of this research study.

4.1.5 Biomechanics Results

Subjective feedback from participants indicated that performing stand-to-sit movements was significantly more challenging than sit-to-stand movements, particularly from a balance control perspective. The experimental and simulated biomechanical data were uploaded to IEEE DataPort and are publicly available for download at https://ieee-dataport.org/documents/measurement-and-simulation-human-sitting-and-standing-movement-biomechanics. The movement durations were time-normalized (0-100%) to allowed for between and within subject averaging. Figure 4.4 shows the hip, knee, and ankle joint angles during stand-to-sit movements from the inverse kinematics. Decreasing joint angles represent hip flexion, knee extension, and ankle dorsiflexion, while increasing joint angles represent hip extension, knee flexion, and ankle plantar flexion. The uncertainties are ± one standard deviation across multiple subjects (n=9) and trials (20 trials/subject), equating to 180 total trials. There were minor variations in the joint kinematics between and within subjects, as demonstrated by the small standard deviations. These simulated joint kinematics both qualitatively and quantitively agree with previous experimental work on robotic leg prostheses for sitting and standing movements [145]–[147].

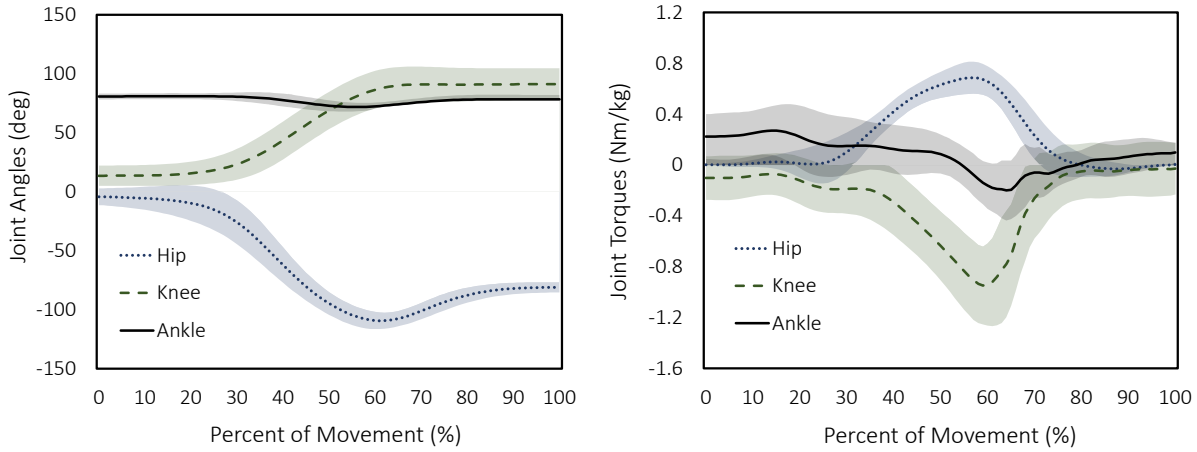

Figure 4.4. The average hip, knee, and ankle joint angles (°) and torques (Nm/kg) during stand-to-sit movements from inverse kinematics and dynamics, respectively. The joint torques are normalized to total body mass and are reported per leg. Subjects were healthy young adults. The uncertainties are ± one standard deviation across multiple subjects (n=9) and trials (20 trials/subject).

Figure 4.4 also shows the calculated hip, knee, and ankle joint torques per leg during stand-to-sit movements from inverse dynamics, normalized to total body mass; the corresponding peak values were 0.7 ± 0.1 Nm/kg, 1.1 ± 0.3 Nm/kg, and 0.4 ± 0.1 Nm/kg, respectively. The uncertainties are ± one standard deviation across multiple subjects and trials. The relatively good quantitatively agreement in peak knee joint torques between the biomechanical simulations (1.1 ± 0.3 Nm/kg) and previous experimental research on robotic leg prostheses and exoskeletons for sitting and standing movements (0.8-1.0 Nm/kg) [138]–[140], [142], [145] supported the model validation. Note that these joint torques are high enough to backdrive the actuators of some robotic leg prostheses and exoskeletons under research and development (e.g., requiring 1-3 Nm of backdrive torque) [81], [83]–[85] and thus theoretically capable of energy regeneration and storage while sitting down, assuming a regenerative motor driver.

Figure 4.5 shows the simulated hip, knee, and ankle joint mechanical powers per leg during stand-to-sit movements and normalized to total body mass. The hip absorbed the largest peak negative mechanical power (1.8 ± 0.5 W/kg), followed by the knee (0.8 ± 0.3 W/kg) and ankle (0.2 ± 0.1 W/kg). Previous experimental work on robotic leg prostheses reported 0.7-0.8 W/kg of peak knee joint mechanical power during sitting and standing movements [140], [143], [145]. The strong quantitative agreement between the simulated (0.8 ± 0.3 W/kg) and experimental (0.7-0.8 W/kg) peak knee joint mechanical powers further supported the model validation. The joint mechanical powers were integrated over time to calculate the joint mechanical energy generated and absorbed while sitting down (i.e., the mechanical energy theoretically available for electrical energy regeneration). The negative mechanical work on the hip, knee, and ankle joints were 0.82

± 0.13 J/kg, 0.35 ± 0.17 J/kg, and 0.06 ± 0.03 J/kg, respectively. It is important to mention that muscle work, not necessarily joint work, is associated with the metabolic energetics of human movements. Accordingly, the design and control of a regenerative actuation system based on joint mechanical work and power alone could bring about a metabolic penalty such that the net joint work is negative but some muscles crossing the joint are doing positive work. This knowledge of the musculoskeletal biomechanics is especially pertinent to exoskeleton devices, which operate in parallel with human muscles.

Although the most negative mechanical work was done on the hip joint while sitting down, and therefore theoretically has the greatest potential for energy regeneration [153], robotic leg prostheses and exoskeletons with regenerative actuators have focused on knee designs [69], [83], [86]–[89], [91], [94], [97]–[102], [104], [106]–[108]; though the robotic ankle prosthesis in [113] regenerated electrical energy during stance phase. The knee joint is typically preferred for energy regeneration given that 1) the hip is more structurally complex (e.g., more degrees of freedom); and 2) many wearable robotic devices, especially prosthetic legs, do not include a hip component. This study on stand-to-sit biomechanics for energy regeneration was published in the IEEE Transactions on Medical Robotics and Bionics [33]. The human joint mechanical energetics from this research study were used in the following section to simulate a robotic lower-limb exoskeleton being backdriven and regenerating electrical energy.

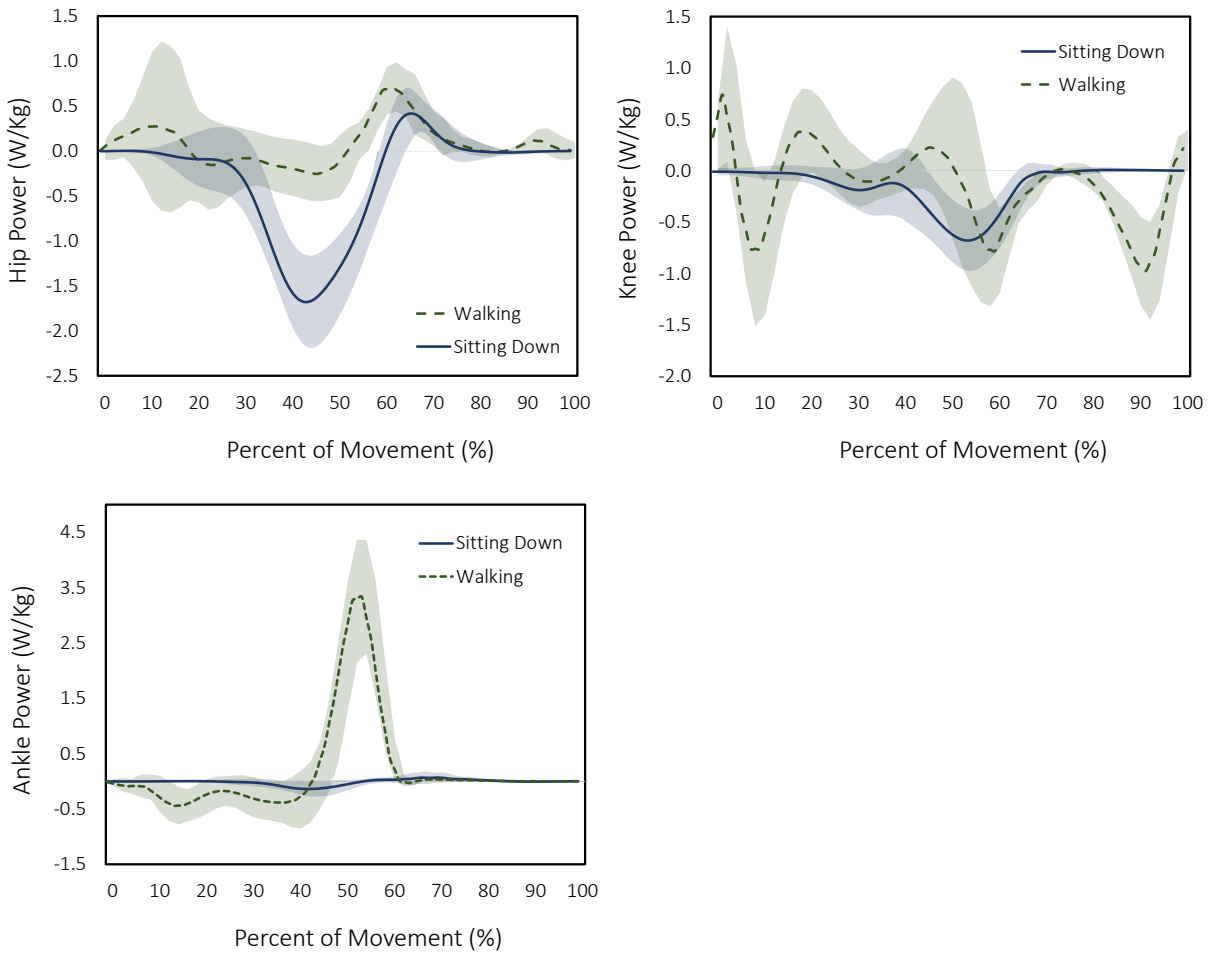

Figure 4.5. The average hip, knee, and ankle joint mechanical power (W/kg) per leg during level-ground walking at natural cadence (n=19) and stand-to-sit movements (n=9), normalized to total body mass. Subjects were all healthy young adults. The positive and negative values represent joint power generation and absorption, respectively. The walking data were taken from Winter [54], the trajectories of which begin and end with heel-strike. The uncertainties are ± one standard deviation across multiple subjects and trials.

4.2 Regenerative Actuation

4.2.1 Introduction

Backdriveable actuators with energy regeneration can extend the battery-powered operating durations and/or decrease the onboard battery weight of robotic leg prostheses and exoskeletons. When backdriven by an external load, the electromagnetic actuator can convert some of the otherwise dissipated mechanical energy into electrical energy while concurrently providing a braking torque to decelerate the load. Building on the biomechanical modelling and simulation work in Section 4.1, here an empirical characterization of a robotic lower-limb exoskeleton was carried out using a joint dynamometer system and an electromechanical motor model in order to 1) calculate the actuator efficiency and 2) simulate electrical energy regeneration and storage during stand-to-sit movements and level-ground walking. Because the motor driver of the exoskeleton does not allow for electrical power flow back to the onboard battery, the benchtop testing was conducted during the standard motoring operation (i.e., electrical-to-mechanical power conversion) and bidirectional symmetry of the actuator efficiency was assumed for the energy regeneration performance calculations.

4.2.2 Robotic Lower-Limb Exoskeleton

A research-grade, robotic lower-limb exoskeleton (Exo-H3, Technaid) was used as an experimental platform to study energy regeneration and storage during stand-to-sit movements and level-ground walking. The device weighs ~12 kg and includes two leg modules, each with a thigh, shank, and foot segment, and a torso, which houses an onboard rechargeable lithium-iron phosphate battery (22 V nominal voltage at 12 Ah capacity) [154]. Six backdriveable actuators, each consisting of a harmonic drive and a wye-wound brushless DC motor, provide bilateral hip, knee, and ankle joint actuation in the sagittal-plane (i.e., six degrees of freedom in total) over the following ranges of motion: 30° plantarflexion and dorsiflexion about the ankle, 105° flexion and 5° extension about the knee, and 105° flexion and 30° extension about the hip. Each motor is rated at 19 V; has a terminal resistance and inductance phase-to-phase of 0.207 Ohms and 0.169 mH, respectively; a torque constant of 0.0375 Nm/A; a speed constant of 225 rpm/V; and a rotor inertia of 0.044 kg-cm^2. The harmonic gearing has a transmission ratio of 160:1 and weighs ~0.24 kg. The final peak torque output of motor-transmission system is 152 Nm with a reflected inertia of 1,126 kg-cm^2 [154].

An external Controller Area Network (CAN) bus was used to communicate with the main onboard controller of the exoskeleton, which runs real-time control algorithms and interacts with the electronic drives of each motorized joint by acquiring sensor feedback and controlling the actuators. Each joint is equipped with an electronic drive board, which performs data acquisition of the onboard sensors, including those for joint angular position, interaction torque, and motor

current (see Figure 4.6); the motor drivers are also embedded into the joint boards. Pressure sensors in the feet provide data regarding the foot-ground contact via center of pressure movement. The main controller and the individual joint controllers communicate at 1 Mbps using a real-time network based on CAN technology [154]; each actuated joint can thus be controlled independently. The exoskeleton uses a hierarchical control architecture. The high-level controller includes individual controllers for different locomotor activities, which are manually selected using a mobile interface, and the low-level controller operates in either position control with prescribed joint kinematics, torque control, or stiffness control, which emulates a virtual spring system. The control architecture is also open, which allows for integration with other devices and/or systems such as neural interfaces or computer vision.

Although the exoskeleton can provide locomotor assistance to individuals with spinal cord injury, the relative backdrivability of the actuated joints could also benefit those with partial motor control (e.g., elderly and/or persons with osteoarthritis or poststroke) by promoting greater user participation and physical rehabilitation [154]. According to the exoskeleton manufacturer, the actuator backdrive torque is ~12 Nm, which is the minimum torque needed to overcome the mechanical impedance (i.e., reflected inertia and friction) to backdrive the motor through its transmission. Backdriveable actuators, either through series elasticity [66]–[71], [73] or low-impedance transmissions [79]–[85], can allow for dynamic physical interactions between the human-exoskeleton-environment system, in addition to energy recycling during periods of negative mechanical work [12]. For robotic exoskeletons with high output impedance, the external loads experienced during daily locomotor activities might be insufficient to overcome the reflected inertia of the actuator and the transmission friction, and thus unable to backdrive the joints and regenerate electrical energy.

4.2.3 Dynamometer Testing

A joint dynamometer (Biodex) was used to measure the mechanical power output of the robotic exoskeleton to calculate the actuator efficiency. These dynamometer systems are commonly used in physical rehabilitation for isolated, single-joint testing to evaluate torque-angle $(\tau - \theta)$ and torque-angular speed $(\tau - \omega)$ relationships [155], from which joint mechanical work and power can be calculated. The left exoskeleton knee was used for testing. Straps were used to secure the exoskeleton leg to the dynamometer attachment and the trunk segment to the seat; this helped prevent relative movement and misalignment between the rotational axes of the exoskeleton knee and the dynamometer shaft, which could cause errors in the torque measurements [155]. Assuming an upright seated posture, the exoskeleton knee was kinematically driven through 12-sinusoidal flexion and extension movements using position control (Figure 4.6). The experiment was repeated four times with 10-minute breaks in between to prevent the exoskeleton motor from overheating (n = 48 total trials). The isokinetic mode of the dynamometer system was selected such that the exoskeleton could drive the shaft throughout the range of motion without

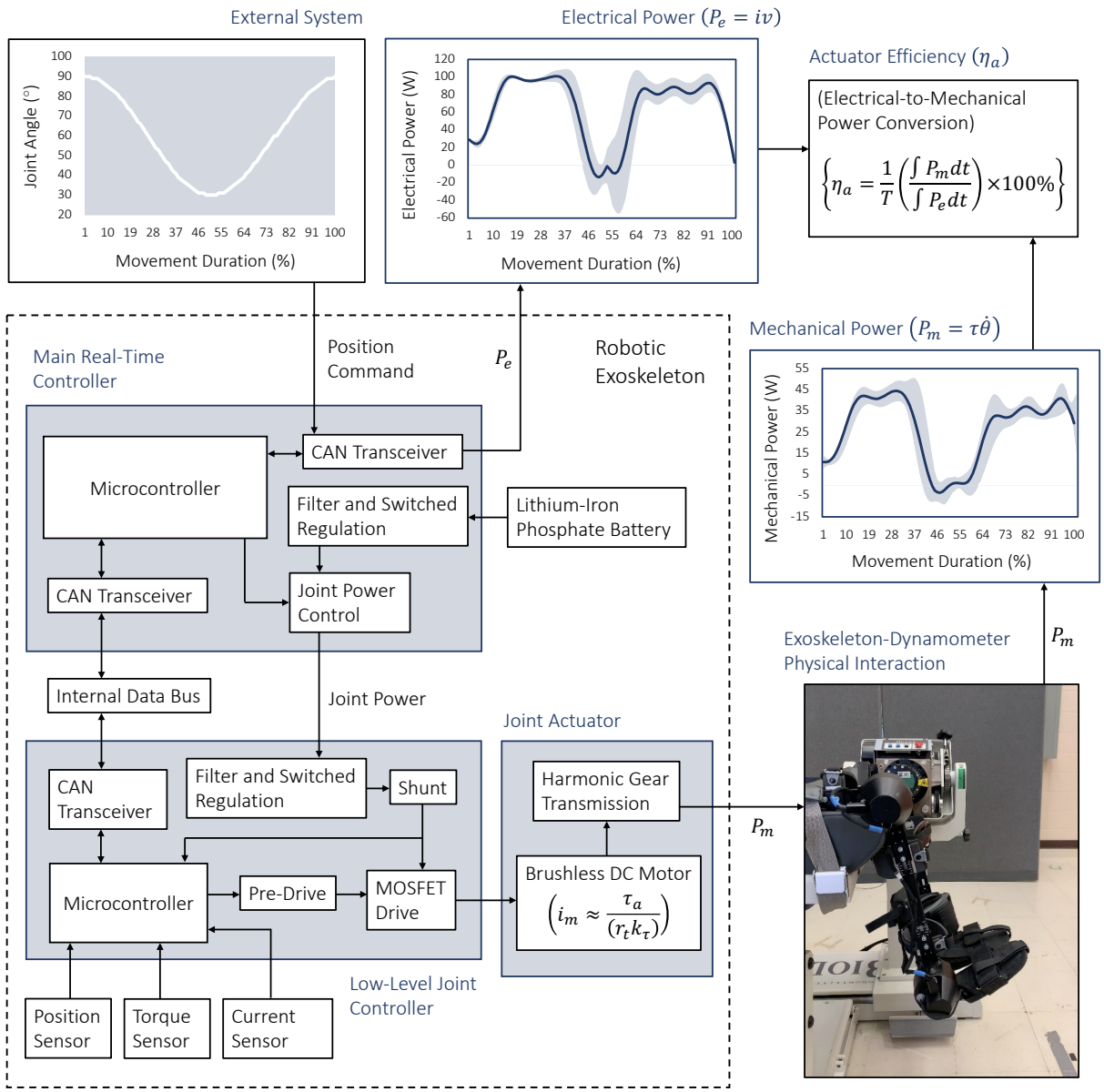

Figure 4.6. Schematic of the exoskeleton benchtop testing using a joint dynamometer system to calculate the actuator efficiency. The exoskeleton knee was kinematically driven through sinusoidal flexion and extension movements while the dynamometer measured the joint torque and angular speed (i.e., the mechanical power output). The onboard exoskeleton sensors, and an electromechanical motor model, were used to estimate the motor voltage and current (i.e., the electrical power input). The actuator efficiency is the ratio of the average instantaneous power outputs to inputs over the steady-state time intervals. The nomenclature are described in the text. Details regarding the exoskeleton control architecture were taken from [145].

interference since the controlled speed threshold of the dynamometer far exceeded that of the exoskeleton joint. Compared to user testing, the benchtop experiments allowed the exoskeleton performance to be quantified without the additional complexities of the human musculoskeletal system in terms of muscle activations and inertial dynamics of the biological leg.

During the benchtop testing, the exoskeleton battery voltage and knee actuator torque were measured by the onboard sensors, and the joint torque and angular speed were measured by the dynamometer with a torque measurement accuracy of ± 1% of full scale (500 ft.lbs). The dynamometer performed an automatic gravity correction on the measured joint torques by measuring the combined exoskeleton-dynamometer segment weight and applying the gravity correction based on the direction of shaft rotation. The weight was taken with the dynamometer such that the leg was positioned horizontally (i.e., maximum knee extension), wherein the gravity effect was highest, and a torque measurement was made. Depending on the exoskeleton leg orientation with respect to gravity during the testing, the weight correction was either added or subtracted to the measured joint torques. Data were sampled at 100 Hz and filtered during postprocessing using a 10th-order low-pass Butterworth filter with an 8 Hz cut-off frequency and normalized to 0-100% of the movement duration. Given that accelerations of the combined exoskeleton-dynamometer segment could produce unwanted inertial loads at the beginning and end of flexion and extension, only the middle portions of each movement, which were relatively constant speed, were used for the data analyses (i.e., 15-35% and 65-85% of the overall sinusoidal movement).

4.2.4 Actuator Efficiency

The actuator efficiency was calculated based on measurements from the dynamometer and the onboard exoskeleton sensors, in addition to an electromechanical motor model. The exoskeleton motor converts electrical power (P_e) to mechanical power (P_m) during motoring operation. The mechanical power output is the product of the joint torque (τ_j) and angular speed $(\dot{\theta}_j)$ and the electrical power input is the product of the motor winding current (i_m) and voltage (v). When backdriven by an external load, the motor can operate like a generator, converting mechanical power to electrical power. The actuator efficiency (η_a) during motoring operation is the ratio of electrical-to-mechanical power conversion $\left\{\eta_a = \frac{1}{T}\left(\frac{\int \tau_j \dot{\theta}_j dt}{\int i_m v_b dt}\right) \times 100\%\right\}$ and vice-versa for energy regeneration when backdriven. The average instantaneous mechanical power output of the exoskeleton actuator was calculated based on the product of the measured joint torque and angular speed by the dynamometer $\{P_m = \tau_j \dot{\theta}_j\}$ at each time step and averaged over the steady-state time intervals. Determining the corresponding average instantaneous electrical power input required some additional consideration, as subsequently discussed.

Many exoskeleton and wearable robotics studies [65]–[71], [73], [83], [88], [89], [93], [96], [102], [105], [113], [156] have used the standard brushed DC electromechanical motor model governed by electrical and mechanical state equations. The motor winding voltage (v_m) can be mathematically modelled by applying Kirchhoff's voltage law:

$$v_m \approx k_\tau \dot{\theta}_m + i_m R_m + L\frac{di_m}{dt} \tag{9}$$

$$\tau_m \approx k_\tau i_m \tag{10}$$

where i_m is the motor winding current, k_τ is the motor torque constant, R_m is the phase resistance of the motor windings, and τ_m and $\dot{\theta}_m$ are the motor torque and angular speed, respectively. The motor inductance $\left(L\frac{di_m}{dt}\right)$ is usually relatively small and thus often omitted [65], [68], [105], [156]. The relationships between the torque-current and speed-voltage data characterize the motor torque constant and back EMF constant, respectively; these constants are device-specific and depend on the motor topology and materials. Although the exoskeleton has onboard motor current sensors, the raw data are not available through the external CAN bus. However, the actuator torque (i.e., the combined motor-transmission system output) is provided. The motor phase current (i_m) was therefore back-calculated according to

$$i_m \approx \frac{\tau_a}{r_t k_\tau} \tag{11}$$

where τ_a is the actuator torque at each time step, r_t is the fixed transmission ratio of the harmonic gearing (160:1), and k_τ is the motor torque constant (0.0375 Nm/A) estimated by the motor manufacturer. The motor winding voltage (v_m) could not be directly solved for using Equation (9) since the onboard sensors do not measure the motor angular speed $\left(\dot{\theta}_m\right)$ and the resolution of the joint angular position data was insufficient to back-calculate $\dot{\theta}_m$ using the transmission ratio (r_t). For simplicity, the measured battery voltage was assumed equal to the voltage between the motor phases. The electrical power input used to drive the exoskeleton knee joint through flexion and extension was estimated by $\{P_e = i_m v_b\}$, where v_b is the measured battery voltage reported by the external CAN bus. Note that the reference axes of the motor variables and parameters were assumed to be equivalent since this information was not provided by the manufacturer. In other words, the three phase windings were assumed to be identical and a single set of parameters was used.

The actuator efficiency (η_a) over the steady-state time intervals was the ratio between the average instantaneous mechanical power output and electrical power input. In theory, the power losses (P_{loss}) are mainly due to Joule heating [156], which is expressed by $\{P_{loss} = i_m^2 R_m\}$. For example, in the MIT Cheetah, 76% of the energy dissipation was attributed to Joule heating [74]. Power losses from battery self-discharging and heating of the motor driver are relatively small and thus often assumed to be negligible [65], [69], [92], [97]. The actuator efficiency across trials

Table 4.1. The average hip, knee, and ankle joint mechanical work (J/kg) per stride in healthy young adults (n=19) walking at natural cadence (110 ±8 steps/min at 1.436 m/s) and normalized to total body mass [17]. "Total Work" is the combined mechanical energies from the hip, knee, and ankle joints and the "Net Joint Work" is the net mechanical work on each joint. Note that 42.6% of the total lower-limb joint mechanical work is negative (i.e., theoretically available for energy regeneration and storage). The results are for one leg.

	Positive Work (J/kg)	Negative Work (J/kg)	Net Joint Work (J/kg)
Hip Joint	0.143	-0.054	0.089
Knee Joint	0.082	-0.247	-0.166
Ankle Joint	0.345	-0.122	0.223
Total Work	0.570	-0.423	0.146

Table 4.2. The average hip, knee, and ankle joint mechanical work (J/kg) per stand-to-sit movement in healthy young adults (~1 repetition per 2.25 seconds) and normalized to total body mass. The averages were taken across multiple subjects (n=9) and trials (20 trials/subject). "Total Work" is the combined mechanical energies from the hip, knee, and ankle joints and the "Net Joint Work" is the net mechanical work on each joint. Note that 91.2% of the total lower-limb joint mechanical work is negative (i.e., theoretically available for energy regeneration and storage). The results are for one leg.

	Positive Work (J/kg)	Negative Work (J/kg)	Net Joint Work (J/kg)
Hip Joint	0.085	-0.780	-0.694
Knee Joint	0.002	-0.331	-0.328
Ankle Joint	0.024	-0.049	-0.026
Total Work	0.111	-1.160	-1.048

(n = 48) was studied to account for any winding temperature changes over time. Although the exoskeleton is somewhat backdriveable, and thus theoretically capable of energy regeneration during negative mechanical work, the motor driver does not allow for electrical power flow back to the onboard battery. Consequently, for the energy regeneration performance calculations, the actuator efficiency was assumed to be bidirectionally symmetric based on the electrical-to-mechanical power conversion efficiency value during the motoring operation. There are ongoing discussions with the exoskeleton manufacturer as to how to safely and effectively regenerate energy without potential damage to the actuator, motor driver, and/or onboard battery.

4.2.5 Regeneration Results

The calculated actuator efficiency from the benchtop testing was ~41%. Assuming an 80-kg user, the robotic exoskeleton could theoretically regenerate 76.1 J of total electrical energy while sitting down since the total negative lower-limb joint mechanical work backdriving the actuators is 185.6 J. Backdriving the same regenerative actuators during walking using the joint biomechanics data by Winter [57], 27.7 J of total electrical energy could theoretically be regenerated per stride since the total negative lower-limb joint mechanical work is 67.7 J. These energy regeneration performance calculations assume 1) regenerative braking from bilateral hip, knee, and ankle

joints; 2) bidirectionally symmetric and constant (i.e., torque and speed independent) actuator efficiencies; 3) power losses only from the motor-transmission system; 4) energy regeneration over the entire negative joint mechanical power range; and 5) identical actuator efficiencies at the hip, knee, and ankle joints. The exoskeleton and human joints were assumed to be kinematically constrained together, therein ignoring any relative translations and rotations. Tables 4.1 and 4.2 summarize the hip, knee, and ankle joint mechanical energies during level-ground walking and stand-to-sit movements, respectively, the quantities of which were used to simulate the robotic exoskeleton performance.

Integrating the positive joint mechanical powers in Figure 4.5 can provide insight into the electrical power consumption (i.e., battery performance) of robotic leg prostheses and exoskeletons during level-ground walking. Here the total positive mechanical work per stride is the sum of the positive mechanical work performed by bilateral hip, knee, and ankle joints. Based on these calculations, human locomotion would require 91.2 J of total positive lower-limb joint mechanical work per stride, which equates to 222.4 J of total electrical energy consumption by the robotic device, assuming an 80-kg bodyweight and the aforementioned actuator efficiency model. Using the parameter values of the lithium-iron phosphate battery of the exoskeleton (22.4 V at 12 Ah for a total battery capacity of 967.7 kJ), the device could theoretically walk 8,701 steps per battery charge. The battery-powered operating duration could therefore be extended by an additional 14.3% (i.e., 1,240 additional steps for 9,941 total steps (T_{steps})) by regenerating energy during walking such that

$$T_{steps} = \frac{C_b}{(E_{-e} - E_{+e})} \times 2 \tag{12}$$

where C_b is the total capacity of the lithium-iron phosphate battery, E_{-e} is the total electrical energy consumption by the robotic exoskeleton per stride, and E_{+e} is the total electrical energy that could be theoretically regenerated per walking stride. Based on Equation (12), the battery-powered operating duration could be extended by an additional 0.5% (i.e., 41 additional steps for 8,742 total steps) by regenerating energy while sitting down, assuming only 60 stand-to-sit movements per day per charge [140]. The positive and negative values represent the electrical power flowing in and out of the battery during braking and motoring operations, respectively. These calculations assume a regenerative motor driver to control the bidirectional flow of electrical power between the motor and onboard battery. Also, the electrical-to-chemical energy conversion efficiency of the battery was ignored.

Although regenerating energy during level-ground walking can produce more electricity than that from stand-to-sit movements per day (i.e., assuming able-bodied biomechanics and activity levels), there are potential benefits to energy regeneration while sitting down. Control of regenerative actuators is notoriously challenging. However, stand-to-sit movements could have higher tolerances to reference tracking errors since the joint mechanical energies are almost

entirely negative (Figure 4.5 and Table 4.2). In contrast, energy regeneration during walking could require more robust actuator control since the joint mechanical energies are intermittent, multidirectional, and time varying (Figure 4.5 and Table 4.1). Inaccurate and/or delayed reference tracking could result in regenerating energy during periods of positive mechanical work. Unlike regenerative braking, generating electricity by exerting positive mechanical work would require the human muscles, especially in exoskeleton applications, to actively backdrive the actuator, which could increase the metabolic power consumption and thus decrease locomotor efficiency [86]–[89]. These energetic consequences would be especially pertinent to aging and rehabilitation populations who already exhibit more inefficient walking energetics [2]. In summary, there are potential advantages and disadvantages to energy regeneration during different locomotor activities of daily living. For optimal efficiency and battery performance, robotic leg prostheses and exoskeletons should be designed to recover some of the otherwise dissipated energy during many different locomotor activities that involve negative mechanical work (e.g., walking, sitting down, and ramp and stair descent).

4.3 Chapter Summary

In this chapter, an energy regeneration system was developed and evaluated using mathematical and computational models of human and wearable robotic systems to simulate electrical energy regeneration and storage during locomotor activities of daily living. Regenerative actuators can improve the energy efficiency and extend the battery-powered operating durations or decrease the weight of the onboard batteries by converting some of the otherwise dissipated energy during negative mechanical work into electrical energy. Although previous studies of robotic leg prostheses and exoskeletons with regenerative actuators have focused on steady-state level-ground walking, persons with mobility impairments (e.g., older adults and/or those with physical disabilities) tend to walk slower and take fewer steps per day, therein limiting the potential for energy regeneration from steady-state level-ground walking. Furthermore, as illustrated by the research in Chapter 3, a relatively small percentage of real-world walking environments consists of steady-state level-ground terrain.

Conversely, sit-to-stand and stand-to-sit movements are common locomotor activities of persons with mobility impairments and are prerequisites to walking. Motivated to explore energy regeneration during other locomotor activities of daily living, Section 4.1 quantified the lower-limb joint mechanical work and power while sitting down using parameter identification and inverse dynamic simulations of subject-specific optimized biomechanical models to estimate the joint mechanical energy theoretically available for electrical energy regeneration and storage. The simulations showed that the hip joint absorbed the largest peak negative mechanical power, followed by the knee and ankle. These results differ from previous studies for other human movements, which showed that the knee joint has the greatest potential for energy regeneration during level-ground walking.

Building on the biomechanical modelling and simulation work in Section 4.1, an empirical characterization of a robotic lower-limb exoskeleton was carried out in Section 4.2 using a joint dynamometer system and an electromechanical motor model. The objectives of this research study were to 1) calculate the actuator efficiency and 2) simulate energy regeneration and storage while sitting down and level-ground walking using the exoskeleton parameters. In other words, the human joint mechanical energetics from Section 4.1 were used to simulate an exoskeleton being backdriven and regenerating energy. Since the motor driver of the physical exoskeleton does not allow for electrical power flow back to the onboard battery, the benchtop testing with the joint dynamometer was conducted during motoring operation (i.e., electrical-to-mechanical power conversion) and bidirectional symmetry of the actuator efficiency was assumed for the energy regeneration performance calculations. The simulation results showed that regenerating energy during stand-to-sit movements provide small improvements in energy efficiency and battery-powered operating durations.

5. CONCLUSIONS AND FUTURE RESEARCH

5.1 Summary of Contributions

This thesis research focused on the development and evaluation of environment classification and energy regeneration systems using mathematical, computational, and machine learning methods to support the energy-efficient actuation and automated control of next-generation robotic leg prostheses and exoskeletons for real-world locomotor assistance. In doing so, the following contributions were made:

Environment Sensing and Classification

- Quantitatively reviewed and compared different statistical pattern recognition and machine learning algorithms used for image classification of walking environments, in addition to the experimental image datasets.
- Developed one of the first environment recognition systems powered by deep learning and computer vision to predict the oncoming walking environments prior to physical interaction, which can allow for more accurate and robust automated high-level control decisions (e.g., locomotion mode transitions).
- Quantitatively reviewed and compared the latest deep convolutional neural networks used for image classification of walking environments.
- Developed the "ExoNet" database - the largest and most diverse open-source dataset of wearable camera images of indoor and outdoor real-world walking environments.
- Introduced a novel hierarchical labelling architecture with 12 individual classes and manually labelled over 923,000 images in the ExoNet database using this new architecture.
- Trained and tested over a dozen state-of-the-art deep convolutional neural networks on ExoNet for image classification and automatic feature engineering, including: EfficientNetB0; InceptionV3; MobileNet; MobileNetV2; VGG16; VGG19; Xception; ResNet50; ResNet101; ResNet152; DenseNet121; DenseNet169; and DenseNet201.
- Quantitatively evaluated and compared the benchmarked CNN architectures and their environment classification predictions using an operational metric called "NetScore", which balances the classification accuracy with the architectural and computational complexities (i.e., important for onboard real-time inference with mobile computing devices). The comparative analyses showed that the EfficientNetB0 network achieved the highest test accuracy; VGG16 the fastest inference time; and MobileNetV2 the best NetScore and least number of parameters and computing operations, which can inform the optimal architecture design or selection depending on the desired system performance.
- Overall, developed an environment sensing and classification system that provides a large-scale benchmark for future research in environment-adaptive locomotor control.

Energy Regeneration

- Reviewed the state-of-the-art in backdriveable actuators with energy regeneration for legged and wearable robotic devices.
- Introduced the importance of energy regeneration and storage during locomotor activities of daily living other than steady-state level-ground walking, which has been the focus of previous research.
- Proposed the application of energy regeneration during stand-to-sit movements such that the otherwise dissipated energy during the periods of negative joint mechanical work can be converted into electrical energy and used to extend the battery-powered operating durations or decrease the weight of the onboard batteries.
- Conducted motion capture experiments with human subjects to quantify the lower-limb joint biomechanics during stand-to-sit movements.
- Custom-designed a lower-limb biomechanical model and generated the dynamic equations of the system by leveraging linear graph theory and symbolic computing.
- Developed and implemented parameter identification and inverse dynamic simulations of subject-specific optimized biomechanical models to calculate the negative joint mechanical work and power while sitting down (i.e., the mechanical energy theoretically available for electrical energy regeneration).
- Created an open-source dataset of experimental and simulated joint biomechanics data during stand-to-stand movements to support the research community.
- Performed benchtop testing of a robotic lower-limb exoskeleton using a joint dynamometer system, which, together with an electromechanical motor model, was used to 1) calculate the actuator efficiency and 2) simulate energy regeneration during stand-to-sit movements and level-ground walking using the exoskeleton parameters.
- Showed that energy regeneration and storage during stand-to-sit movements provide small improvements in energy efficiency and battery-powered operating durations.

In addition to robotic leg prostheses and exoskeletons, these principles of energy regeneration and environment classification could extend to humanoids, autonomous legged robots, powered wheelchairs, and other mobility assistive technologies.

5.2 Recommendations for Future Research

5.2.1 Environment Recognition

Although this thesis research helped advance the state-of-the-art in image classification of walking environments via the applications of deep learning and computer vision, there are still many promising and relatively untapped areas of research that could support the development of next-generation environment-adaptive locomotor control systems for robotic leg prostheses and exoskeletons. Some potential directions for future research include: 1) the application of depth cameras for 3D environment sensing; 2) the classification of sequential walking images using recurrent neural networks; and 3) the fusion of environment data with mechanical, inertial, and/or neuromuscular data for automated high-level control and decision making.

5.2.1.1 Depth Sensing

A potential limitation of the ExoNet database developed in Chapter 3 is the 2D nature of the environment information. Many researchers have similarly used a wearable RGB camera for passive environment sensing [36]–[43] (see Table 2.1). Although multiple RGB cameras could be used to capture 3D environment information, comparable to how the human visual system uses triangulation for depth perception [30], each pixel in an RGB image contains only light intensity information. Several researchers have explored using depth cameras to explicitly capture images containing both light intensity information and distance measurements [29], [31], [44]–[50]. These range imaging systems work by emitting infrared light and measuring the light time-of-flight between the camera and oncoming walking environment to calculate distance. Depth sensing can uniquely extract environmental features like step height and width, which can improve the mid-level control of robotic leg prostheses and exoskeletons (e.g., increasing the actuator joint torques to assist with steeper stairs).

Despite the aforementioned benefits, depth measurement accuracy typically degrades in outdoor lighting conditions (e.g., sunlight) and with increasing measurement distance [157], [158]. Consequently, most environment recognition systems using depth cameras have been tested in controlled laboratory environments and/or have had limited capture volumes with 1-2 m of maximum range imaging [31], [44]–[47]. These systems also require an onboard accelerometer or IMU to transform the 3D environment information from the camera coordinate system into global coordinates [29], [31], [46], [48]–[50]. Assuming mobile computing (i.e., untethered and no wireless communication to cloud computing), the use of depth cameras for active environment sensing could potentially require robotic leg prostheses and exoskeletons to have onboard microcontrollers with high computing power and low power consumption [44]. In such a case, the current embedded systems would need significant modifications to support the real-time

processing and classification of depth images. These practical limitations motivated the decision to use RGB images for the environment sensing and classification.

5.2.1.2 Recurrent Neural Networks

The comparative analysis in Chapter 3 showed that the EfficientNetB0 network [122] achieved the highest image classification accuracy on the ExoNet database during inference (73.2%). However, for environment-adaptive control of robotic leg prostheses and exoskeletons, higher classification accuracies would be desired given that misclassifications could cause loss-of-balance and injury [159]. Although the convolutional neural networks that were tested are considered state-of-the-art in computer vision, the architectures included only feedforward connections such that the walking environments were independently classified frame-by-frame without knowledge of the preceding decisions. Sequential data over time could potentially improve the image classification accuracy and robustness, especially during steady state environments. This technique is analogous to how light-sensitive receptors in the human eye capture dynamic images to control locomotion, known as optical flow [30]. Sequential data could be classified using majority voting [24], [44], [45] or Transformers [160] or recurrent neural networks (RNNs) [43], [50], [51]. Majority voting stores sequential decisions in a vector and generates a classification prediction based on the majority of stored decisions. These systems can attenuate misclassifications by filtering over the decision vector, thereby decreasing the likelihood of a fall resulting from an incorrect high-level control decision.

In comparison, recurrent neural networks process sequential inputs while maintaining an internal hidden state vector that implicitly contains temporal information. However, training RNNs can be challenging due to exploding and vanishing gradients. Although these networks were designed to learn long-term dependencies, theoretical and empirical evidence has shown that they struggle with storing sequential information over long periods [55], [161]. To mitigate this issue, RNNs can be supplemented with an explicit memory module like a neural Turing machine or long short-term memory (LSTM) network, which can improve gradient flow. Fu's research group [50] explored the use of sequential data for environment classification. Sequential decisions from a baseline CNN were fused and classified using a recurrent neural network, LSTM network, majority voting, and a hidden Markov model (HMM). The baseline network achieved 92.8% classification accuracy across five environment classes. Supplementing the baseline CNN with the RNN, LSTM network, majority voting, and HMM resulted in 96.5%, 96.4%, 95%, and 96.8% image classification accuracies, respectively [50]. Although sequential data could improve the image classification accuracy of walking environments, and therefore increase the robustness of the automated locomotor control decisions, this technique often requires longer computation times, which could impede real-time implementation.

5.2.1.3 Sensor Fusion

Given that information about the oncoming walking environment does not explicitly represent the user's locomotor intent, though it constrains the movement possibilities, data from computer vision should supplement, rather than replace, the automated locomotion mode control decisions based on user-dependent measurements like mechanical, inertial, and/or neuromuscular data. The images from the wearable smartphone camera system used in Chapter 3 could be fused with its onboard IMU measurements to improve the high-level control performance. For example, when an exoskeleton or prosthesis user wants to sit down, the acceleration data from the IMU sensors would indicate a stand-to-sit movement rather than level-ground walking, despite level-ground terrain being accuracy detected within the visual field-of-view, as illustrated in the bottom right image of Figure 3.7.

Inspired by previous work [24], [34], [36]–[38], the smartphone IMU data could also help minimize the onboard computational and memory storage requirements via sampling rate control (i.e., providing an automatic triggering mechanism for the image capture). Whereas fast walking speeds could benefit from high sampling rates for continuous classification, standing still does not necessarily require environment information and thus the smartphone camera could be powered-down or the sampling rate decreased to conserve the onboard operating resources. Relatively few researchers have fused environment data with mechanical and/or inertial measurements for automated locomotion mode recognition [23], [24], [26], [28], [29], [34], [47] and only one study [31] has used such information for online environment-adaptive control of a robotic leg prosthesis during walking (i.e., stepping over an obstacle). These limitations in systems integration offer exciting challenges and opportunities for future research such that the optimal multi-sensor data fusion for environment-adaptive control of robotic leg prostheses and exoskeletons remains to be determined.

5.2.2. Energy Regeneration

With regards to energetics, this thesis research explored energy regeneration and storage during locomotor activities of daily living, specifically stand-to-sit movements. Building on the research presented in Chapter 4, some potential directions for future research include: 1) the application of ultracapacitors for energy storage with high power density; 2) the optimization of actuator design parameters for efficiency; and 3) the study of older adults and/or persons with physical disabilities to evaluate the potential for energy regeneration. Note that each of these proposed research areas could benefit from an integrated system model of the human interacting with a wearable robotic device [159]. For instance, a human-exoskeleton system model could provide a computational framework for efficient model-based testing and optimization of different actuator design parameters and/or energy storage devices while taking into consideration changes in the human biomechanics and motor control. This differs from the work in Chapter 4, which developed and used independent mathematical and computational models of the human and exoskeleton systems without accounting for the closed-loop interactions.

5.2.2.1 Ultracapacitors

An interesting topic for future research is to evaluate different onboard energy storage devices for robotic leg prostheses and exoskeletons with regenerative actuators. The energy regeneration performance calculations presented in Chapter 4 were based on the parameter values of the rechargeable lithium-iron phosphate battery of the exoskeleton being tested. Most robotic leg prostheses and exoskeletons use lithium-polymer or lithium-ion batteries for energy storage [8], [32]. Generally speaking, rechargeable batteries are designed with high energy density (e.g., ~100 Wh/kg), which allow for extended operation, but low power density (e.g., 0.1-1 kW/kg), which produce slow charge and discharge rates [162]. In many mechatronic applications, however, the rate at which mechanical energy should be converted into electrical energy for regenerative braking is higher than the rate at which most batteries can absorb energy [94]. In other words, rechargeable batteries tend to have insufficient power densities. These limitations have led some researchers to explore the use of ultracapacitors for energy regeneration and storage in semi-powered leg prostheses [95], [97], [98], [101], [102], [108], [113].

Compared to batteries, ultracapacitors have high power density (e.g., ~10 kW/kg) but low energy density (e.g., 1-10 Wh/kg) [162]. They provide an efficient means of energy storage that is lightweight and inexpensive; can charge and discharge at high rates without damage; and have almost infinite lifecycles [162]. The dynamic and efficient bidirectional power flow characteristics of ultracapacitors make them ideal for applications requiring many fast charge and discharge cycles and avoid the need for dissipative elements like resistors during regenerative braking. The pioneering work by Flowers' lab at MIT in the 1980s used conventional electrolytic capacitors for energy regeneration and storage in their prosthetic leg design [163]–[165] but concluded that the

capacitors at the time were insufficient and needed further development. Ultracapacitors bridge the gap between conventional capacitors and batteries. Although the total energy stored per unit mass in ultracapacitors is usually much smaller than rechargeable batteries, recent breakthroughs in nanotechnology are enabling the fabrication of graphene-based ultracapacitors, which have reached energy densities of 64 Wh/kg [96]. Moving forward, an optimal energy storage system for robotic leg prostheses and exoskeletons with regenerative actuators could include an ultracapacitor, for fast charging and discharging rates, and a battery, for extended operation.

5.2.2.2 Actuator Design Optimization

One of the biggest limitations to energy regeneration is the relatively low efficiencies of traditional motor-transmission systems [15]. The calculated actuator efficiency of the robotic lower-limb exoskeleton in Chapter 4 was ~41% during forward operation (i.e., electrical-to-mechanical power conversion); bidirectional symmetry of the actuator efficiency was assumed to simulate energy regeneration when backdriven (i.e., mechanical-to-electrical power conversion). For comparison, the MIT Cheetah robot has reported one of the highest energy regeneration efficiencies during backdrive operation (~63%) [74]. Two of the main sources of energy losses in robotic actuators, which can characterize the inefficiencies of the system, are Joule heating in the motor windings and friction in the transmission [11], [12].

High transmission ratios can reduce the motor torque needed for legged locomotion, and thus decrease the winding current and associated Joule heating losses. However, high gearing systems tend to increase weight, friction, and reflected inertia, which increase the impedance and reduces backdrivability and the potential for energy regeneration [12]. Alternatively, high torque-density motors can decrease the needed transmission ratios by generating a high output torque, therein circumventing the inefficiencies of high gearing systems, though at the expense of more winding current and thus higher Joule heating losses [11]. For example, Joule heating accounted for ~76% of the energy losses in the MIT Cheetah, which was designed using high torque motors with minimal gearing (i.e., 6:1 transmission ratio) [74].

An open challenge for the research community is to optimize the tradeoff between the actuator output torque and backdrive torque in robotic leg prostheses and exoskeletons in terms of energy efficiency and performance. Many design parameters can effect this tradeoff (e.g., the transmission ratio and efficiency, and the motor terminal resistance, and the torque and speed constants) [11]. Given the complex interactions between these parameters, determining the optimal actuator design through experimental trial-and-error would be difficult. Modelling and simulation could be used to automatically co-optimize the motor and transmission system design parameters to maximize both energy efficiency (including energy regeneration) and performance. This computational framework could also help address some of the simplifying assumptions made in the energy regeneration performance calculations in Chapter 4 (e.g., bidirectional symmetry).

5.2.2.3 Clinical Evaluations

Future research could also study the potential for energy regeneration in individuals with mobility impairments. The energy regeneration performance calculations in Chapter 4 were based on data from healthy young adults, therein requiring several assumptions and extrapolations to aging and rehabilitation populations. Young adults typically walk at 1.4 m/s and take 6,000-13,000 steps/day [2]. However, older adults tend to walk slower and shorter distances. For example, 50% of persons over 65 years age walk fewer than 5000 steps/day [2]. These activity levels are further diminished in individuals with physical disabilities (e.g., persons with incomplete spinal cord injury walk ~1,640 steps/day) [2]. The biomechanics of persons with mobility impairments also differ from healthy young adults. Table 5.1 and Figure 5.1 show the hip, knee, and ankle joint mechanical work and power per stride in older adults walking at preferred speed (1.28 m/s) [57]. The peak power absorption by the knee flexors during late swing is lower in older adults (0.87 W/kg) than young adults (0.98 W/kg), which decrease the leg deceleration prior to heel strike and thus the potential for energy regeneration. Moreover, the peak power generation by the ankle plantarflexors during push-off is lower in older adults (2.48 W/kg) compared to young adults (3.23 W/kg), which could be attributed to muscle weakening and explain the slower walking speeds [57].

These population differences, especially in walking speed, have implications on the energy regeneration performance. Experimental studies of robotic leg prostheses and exoskeletons with regenerative actuators have consistently shown a positive relationship between walking speed and both energy regeneration and efficiency such that faster walking generates more electricity and more efficiently [83], [88]–[90], [104], [106], [107], [113]. For a given back EMF constant, an electromagnetic motor generates a voltage proportional to its rotational speed. Slower walking would backdrive the actuator with lower rotational speeds and therefore generate less electricity (i.e., although for a longer time for the same distance). Motors are also generally less efficient when generating torques at low speeds due to Joule heating losses.

A recent study by Gregg's research lab showed that increasing walking speed with a robotic knee-ankle prosthesis from 0.9 m/s to 1.6 m/s increased the power conversion efficiency of the actuator from 40% to 59% [83]. Feng et al. [113] showed that the ratio of energy regeneration to total power consumption increased from 27% to 35% when walking speed increased from 0.7 m/s to 1.3 m/s with a robotic ankle prosthesis. For battery recharging at slower walking speeds, the motor could be designed or selected with a higher speed constant, or the number of batteries decreased to lower the charging potential. Given the biomechanical and activity level differences between healthy young adults and individuals with mobility impairments, and the implications of such differences on the energy regeneration and efficiency, future research should study older adults and/or those with physical disabilities to improve the energy regeneration performance calculations presented in this thesis.

Table 5.1. The average hip, knee, and ankle joint mechanical work (J/kg) per stride in older adults (n=18) walking at natural cadence (111 ±8 steps/min at 1.28 m/s) and normalized to total body mass [57]. "Total Work" is the combined mechanical energies from the hip, knee, and ankle joints and the "Net Joint Work" is the net mechanical work on each joint. Note that 47.5% of the total lower-limb joint mechanical work is negative (i.e., theoretically available for energy regeneration and storage). The results are for one leg.

	Positive Work (J/kg)	Negative Work (J/kg)	Net Joint Work (J/kg)
Hip Joint	0.223	-0.066	0.157
Knee Joint	0.063	-0.260	-0.197
Ankle Joint	0.225	-0.136	0.089
Total Work	0.511	-0.462	0.049

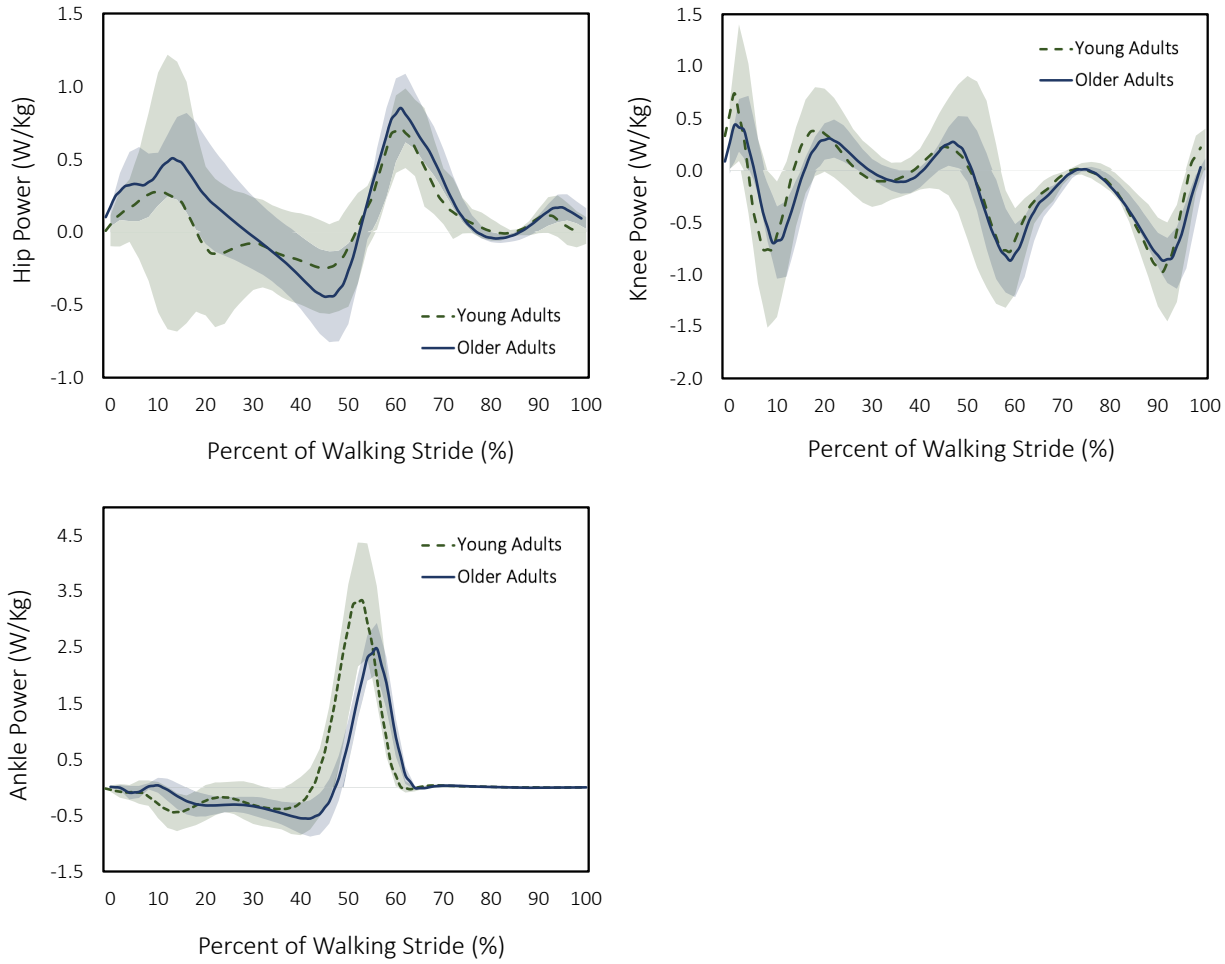

Figure 5.1. The average hip, knee, and ankle joint mechanical power (W/kg) per leg while walking at natural cadence in young adults (n=19) and older adults (n=18), normalized to total body mass. The positive and negative values represent joint power generation and absorption, respectively. Data were taken from Winter [57], the trajectories of which begin and end with heel-strike. The uncertainties are ± one standard deviation across multiple subjects and trials.

REFERENCES

[1] K. Ziegler-Graham, E. J. MacKenzie, P. L. Ephraim, T. G. Travison, and R. Brookmeyer, "Estimating the Prevalence of Limb Loss in the United States: 2005 to 2050," *Arch. Phys. Med. Rehabil.*, vol. 89, no. 3, pp. 422–429, Mar. 2008, doi: 10.1016/j.apmr.2007.11.005.

[2] M. Grimmer, R. Riener, C. J. Walsh, and A. Seyfarth, "Mobility Related Physical and Functional Losses due to Aging and Disease - A Motivation for Lower Limb Exoskeletons," *J. NeuroEngineering Rehabil.*, vol. 16, no. 1, p. 2, Jan. 2019, doi: 10.1186/s12984-018-0458-8.

[3] S. Negrini, J.-A. Mills, C. Arienti, C. Kiekens, and A. Cieza, "'Rehabilitation Research Framework for Patients With COVID-19' Defined by Cochrane Rehabilitation and the World Health Organization Rehabilitation Programme," *Arch. Phys. Med. Rehabil.*, p. S0003999321002240, Mar. 2021, doi: 10.1016/j.apmr.2021.02.018.

[4] M. R. Tucker *et al.*, "Control Strategies for Active Lower Extremity Prosthetics and Orthotics: A Review," *J. NeuroEngineering Rehabil.*, vol. 12, no. 1, p. 1, 2015, doi: 10.1186/1743-0003-12-1.

[5] A. S. Voloshina and S. H. Collins, "Lower Limb Active Prosthetic Systems - Overview," in *Wearable Robotics: Systems and Applications*, Elsevier, 2020, pp. 469–486. doi: 10.1016/B978-0-12-814659-0.00023-0.

[6] D. S. Pieringer, M. Grimmer, M. F. Russold, and R. Riener, "Review of the Actuators of Active Knee Prostheses and their Target Design Outputs for Activities of Daily Living," in *2017 International Conference on Rehabilitation Robotics (ICORR)*, London, Jul. 2017, pp. 1246–1253. doi: 10.1109/ICORR.2017.8009420.

[7] A. M. Dollar and H. Herr, "Lower Extremity Exoskeletons and Active Orthoses: Challenges and State-of-the-Art," *IEEE Trans. Robot.*, vol. 24, no. 1, pp. 144–158, Feb. 2008, doi: 10.1109/TRO.2008.915453.

[8] A. J. Young and D. P. Ferris, "State of the Art and Future Directions for Lower Limb Robotic Exoskeletons," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 2, pp. 171–182, Feb. 2017, doi: 10.1109/TNSRE.2016.2521160.

[9] A. B. Zoss, H. Kazerooni, and A. Chu, "Biomechanical Design of the Berkeley Lower Extremity Exoskeleton (BLEEX)," *IEEEASME Trans. Mechatron.*, vol. 11, no. 2, pp. 128–138, Apr. 2006, doi: 10.1109/TMECH.2006.871087.

[10] R. J. Farris, H. A. Quintero, and M. Goldfarb, "Preliminary Evaluation of a Powered Lower Limb Orthosis to Aid Walking in Paraplegic Individuals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 6, pp. 652–659, Dec. 2011, doi: 10.1109/TNSRE.2011.2163083.

[11] J. Hollerbach, I. Hunter, and J. Ballantyne, "A Comparative Analysis of Actuator Technologies for Robotics," in *The Robotics Review 2*, 1992, pp. 299–342.

[12] P. L. García, S. Crispel, E. Saerens, T. Verstraten, and D. Lefeber, "Compact Gearboxes for Modern Robotics: A Review," *Front. Robot. AI*, vol. 7, p. 103, Aug. 2020, doi: 10.3389/frobt.2020.00103.

[13] K. Karacan, J. T. Meyer, H. I. Bozma, R. Gassert, and E. Samur, “An Environment Recognition and Parameterization System for Shared-Control of a Powered Lower-Limb Exoskeleton,” in *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, New York City, NY, USA, Nov. 2020, pp. 623–628. doi: 10.1109/BioRob49111.2020.9224407.

[14] B. Laschowski and J. Andrysek, “Electromechanical Design of Robotic Transfemoral Prostheses,” in *ASME 2018 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE)*, Quebec City, Quebec, Canada, Nov. 2018, p. V05AT07A054. doi: 10.1115/DETC2018-85234.

[15] N. Kashiri *et al.*, “An Overview on Principles for Energy Efficient Robot Locomotion,” *Front. Robot. AI*, vol. 5, p. 129, Dec. 2018, doi: 10.3389/frobt.2018.00129.

[16] D. Farina *et al.*, “Toward Higher-Performance Bionic Limbs for Wider Clinical Use,” *Nat. Biomed. Eng.*, May 2021, doi: 10.1038/s41551-021-00732-x.

[17] A. Maryniak, B. Laschowski, and J. Andrysek, “Technical Overview of Osseointegrated Transfemoral Prostheses: Orthopedic Surgery and Implant Design Centered,” *ASME J. Eng. Sci. Med. Diagn. Ther.*, vol. 1, no. 2, p. 020801, May 2018, doi: 10.1115/1.4039105.

[18] H. A. Varol, F. Sup, and M. Goldfarb, “Multiclass Real-Time Intent Recognition of a Powered Lower Limb Prosthesis,” *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 542–551, Mar. 2010, doi: 10.1109/TBME.2009.2034734.

[19] H. Huang, T. A. Kuiken, and R. D. Lipschutz, “A Strategy for Identifying Locomotion Modes Using Surface Electromyography,” *IEEE Trans. Biomed. Eng.*, vol. 56, no. 1, pp. 65–73, Jan. 2009, doi: 10.1109/TBME.2008.2003293.

[20] A. Nasr, B. Laschowski, and J. McPhee, “Myoelectric Control of Robotic Leg Prostheses and Exoskeletons: A Review,” in *ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Virtual, Aug. 2021, vol. 8A, p. 8. doi: 10.1115/DETC2021-69203.

[21] A. Fleming, N. Stafford, S. Huang, X. Hu, D. P. Ferris, and H. (Helen) Huang, “Myoelectric Control of Robotic Lower Limb Prostheses: A Review of Electromyography Interfaces, Control Paradigms, Challenges and Future Directions,” *J. Neural Eng.*, Jul. 2021, doi: 10.1088/1741-2552/ac1176.

[22] M. A. Lebedev and M. A. L. Nicolelis, “Brain-Machine Interfaces: From Basic Science to Neuroprostheses and Neurorehabilitation,” *Physiol. Rev.*, vol. 97, no. 2, pp. 767–837, Apr. 2017, doi: 10.1152/physrev.00027.2016.

[23] L. Du, F. Zhang, M. Liu, and H. Huang, “Toward Design of an Environment-Aware Adaptive Locomotion-Mode-Recognition System,” *IEEE Trans. Biomed. Eng.*, vol. 59, no. 10, pp. 2716–2725, Oct. 2012, doi: 10.1109/TBME.2012.2208641.

[24] D. Wang, L. Du, and H. Huang, “Terrain Recognition Improves the Performance of Neural-Machine Interface for Locomotion Mode Recognition,” in *2013 International Conference on Computing, Networking and Communications (ICNC)*, San Diego, CA, Jan. 2013, pp. 87–91. doi: 10.1109/ICCNC.2013.6504059.

[25] J. A. Spanias, A. M. Simon, K. A. Ingraham, and L. J. Hargrove, “Effect of Additional Mechanical Sensor Data on an EMG-Based Pattern Recognition System for a Powered Leg Prosthesis,” in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, Montpellier, France, Apr. 2015, pp. 639–642. doi: 10.1109/NER.2015.7146704.

[26] M. Liu, D. Wang, and H. Huang, “Development of an Environment-Aware Locomotion Mode Recognition System for Powered Lower Limb Prostheses,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 4, pp. 434–443, Apr. 2016, doi: 10.1109/TNSRE.2015.2420539.

[27] H. Huang, F. Zhang, L. J. Hargrove, Z. Dou, D. R. Rogers, and K. B. Englehart, “Continuous Locomotion-Mode Identification for Prosthetic Legs Based on Neuromuscular–Mechanical Fusion,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2867–2875, Oct. 2011, doi: 10.1109/TBME.2011.2161671.

[28] H. Huang, Z. Dou, F. Zhang, and M. J. Nunnery, “Improving the Performance of a Neural-Machine Interface for Artificial Legs using Prior Knowledge of Walking Environment,” in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Boston, MA, USA, Aug. 2011, pp. 4255–4258. doi: 10.1109/IEMBS.2011.6091056.

[29] N. E. Krausz and L. J. Hargrove, “Sensor Fusion of Vision, Kinetics and Kinematics for Forward Prediction During Walking with a Transfemoral Prosthesis,” *IEEE Trans. Med. Robot. Bionics*, pp. 1–1, 2021, doi: 10.1109/TMRB.2021.3082206.

[30] A. E. Patla, “Understanding the Roles of Vision in the Control of Human Locomotion,” *Gait Posture*, vol. 5, no. 1, pp. 54–69, Feb. 1997, doi: 10.1016/S0966-6362(96)01109-5.

[31] K. Zhang *et al.*, “A Subvision System for Enhancing the Environmental Adaptability of the Powered Transfemoral Prosthesis,” *IEEE Trans. Cybern.*, vol. 51, no. 6, pp. 3285–3297, Jun. 2021, doi: 10.1109/TCYB.2020.2978216.

[32] B. Laschowski, J. McPhee, and J. Andrysek, “Lower-Limb Prostheses and Exoskeletons With Energy Regeneration: Mechatronic Design and Optimization Review,” *ASME J. Mech. Robot.*, vol. 11, no. 4, p. 040801, Aug. 2019, doi: 10.1115/1.4043460.

[33] B. Laschowski, R. S. Razavian, and J. McPhee, “Simulation of Stand-to-Sit Biomechanics for Robotic Exoskeletons and Prostheses With Energy Regeneration,” *IEEE Trans. Med. Robot. Bionics*, vol. 3, no. 2, pp. 455–462, May 2021, doi: 10.1109/TMRB.2021.3058323.

[34] F. Zhang, Z. Fang, M. Liu, and H. Huang, “Preliminary Design of a Terrain Recognition System,” in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Boston, MA, Aug. 2011, pp. 5452–5455. doi: 10.1109/IEMBS.2011.6091391.

[35] B. Kleiner, N. Ziegenspeck, R. Stolyarov, H. Herr, U. Schneider, and A. Verl, “A Radar-Based Terrain Mapping Approach for Stair Detection Towards Enhanced Prosthetic Foot Control,” in *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*, Enschede, Netherlands, Aug. 2018, pp. 105–110. doi: 10.1109/BIOROB.2018.8487722.

[36] R. L. da Silva, N. Starliper, B. Zhong, H. H. Huang, and E. Lobaton, “Evaluation of Embedded Platforms for Lower Limb Prosthesis with Visual Sensing Capabilities,” *arXiv*, Jun. 2020, doi: arXiv:2006.15224.

[37] J. P. Diaz, R. L. da Silva, B. Zhong, H. H. Huang, and E. Lobaton, “Visual Terrain Identification and Surface Inclination Estimation for Improving Human Locomotion with a Lower-Limb Prosthetic,” in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, Jul. 2018, pp. 1817–1820. doi: 10.1109/EMBC.2018.8512614.

[38] G. Khademi and D. Simon, “Convolutional Neural Networks for Environmentally Aware Locomotion Mode Recognition of Lower-Limb Amputees,” in *ASME Dynamic Systems and Control Conference (DSCC)*, Park City, Utah, USA, Oct. 2019, p. 11. doi: 10.1115/DSCC2019-9180.

[39] N. E. Krausz and L. J. Hargrove, “Recognition of Ascending Stairs from 2D Images for Control of Powered Lower Limb Prostheses,” in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, Montpellier, France, Apr. 2015, pp. 615–618. doi: 10.1109/NER.2015.7146698.

[40] B. Laschowski, W. McNally, A. Wong, and J. McPhee, “Preliminary Design of an Environment Recognition System for Controlling Robotic Lower-Limb Prostheses and Exoskeletons,” in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, Toronto, ON, Canada, Jun. 2019, pp. 868–873. doi: 10.1109/ICORR.2019.8779540.

[41] B. Laschowski, W. McNally, A. Wong, and J. McPhee, “ExoNet Database: Wearable Camera Images of Human Locomotion Environments,” *Front. Robot. AI*, vol. 7, p. 562061, Dec. 2020, doi: 10.3389/frobt.2020.562061.

[42] L. Novo-Torres, J.-P. Ramirez-Paredes, and D. J. Villarreal, “Obstacle Recognition using Computer Vision and Convolutional Neural Networks for Powered Prosthetic Leg Applications,” in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, Jul. 2019, pp. 3360–3363. doi: 10.1109/EMBC.2019.8857420.

[43] B. Zhong, R. L. da Silva, M. Li, H. Huang, and E. Lobaton, “Environmental Context Prediction for Lower Limb Prostheses With Uncertainty Quantification,” *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 2, pp. 458–470, Apr. 2021, doi: 10.1109/TASE.2020.2993399.

[44] Y. Massalin, M. Abdrakhmanova, and H. A. Varol, “User-Independent Intent Recognition for Lower Limb Prostheses Using Depth Sensing,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 8, pp. 1759–1770, Aug. 2018, doi: 10.1109/TBME.2017.2776157.

[45] H. A. Varol and Y. Massalin, “A Feasibility Study of Depth Image Based Intent Recognition for Lower Limb Prostheses,” in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, USA, Aug. 2016, pp. 5055–5058. doi: 10.1109/EMBC.2016.7591863.

[46] N. E. Krausz, T. Lenzi, and L. J. Hargrove, "Depth Sensing for Improved Control of Lower Limb Prostheses," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 11, pp. 2576–2587, Nov. 2015, doi: 10.1109/TBME.2015.2448457.

[47] N. E. Krausz, B. H. Hu, and L. J. Hargrove, "Subject- and Environment-Based Sensor Variability for Wearable Lower-Limb Assistive Devices," *Sensors*, vol. 19, no. 22, p. 4887, Nov. 2019, doi: 10.3390/s19224887.

[48] K. Zhang, J. Wang, and C. Fu, "Directional PointNet: 3D Environmental Classification for Wearable Robotics," *arXiv*, p. 8, Mar. 2019, doi: arXiv:1903.06846.

[49] K. Zhang *et al.*, "Environmental Features Recognition for Lower Limb Prostheses Toward Predictive Walking," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 3, pp. 465–476, Mar. 2019, doi: 10.1109/TNSRE.2019.2895221.

[50] K. Zhang, W. Zhang, W. Xiao, H. Liu, C. W. De Silva, and C. Fu, "Sequential Decision Fusion for Environmental Classification in Assistive Walking," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 9, pp. 1780–1790, Sep. 2019, doi: 10.1109/TNSRE.2019.2935765.

[51] B. Zhong, R. L. da Silva, M. Tran, H. Huang, and E. Lobaton, "Efficient Environmental Context Prediction for Lower Limb Prostheses," *IEEE Trans. Syst. Man Cybern. Syst.*, pp. 1–15, 2021, doi: 10.1109/TSMC.2021.3084036.

[52] M. Nouredanesh, A. McCormick, S. Kukreja, and J. Tung, "Wearable Vision Detection of Environmental Fall Risk using Gabor Barcodes," presented at the IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob), Singapore, Jun. 2016. doi: 10.1109/BIOROB.2016.7523751.

[53] M. Nouredanesh, A. McCormick, S. Kukreja, and J. Y. Tung, "Wearable Vision Detection of Environmental Fall Risks Using Convolutional Neural Networks," *arXiv*, p. 1, Nov. 2016, doi: arXiv:1611.00684.

[54] V. Rai, D. Boe, and E. Rombokas, "Vision for Prosthesis Control Using Unsupervised Labeling of Training Data," in *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)*, Munich, Germany, Jul. 2021, pp. 326–333. doi: 10.1109/HUMANOIDS47582.2021.9555789.

[55] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[56] B. Laschowski, W. McNally, A. Wong, and J. McPhee, "Comparative Analysis of Environment Recognition Systems for Control of Lower-Limb Exoskeletons and Prostheses," in *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, New York City, NY, USA, Nov. 2020, pp. 581–586. doi: 10.1109/BioRob49111.2020.9224364.

[57] D. A. Winter, *The Biomechanics and Motor Control of Human Gait: Normal, Elderly and Pathological*, 2nd Edition. Waterloo, Canada: Waterloo Biomechanics, 1991.

[58] D. A. Winter, "Energy Generation and Absorption at the Ankle and Knee during Fast, Natural, and Slow Cadences," *Clin. Orthop.*, no. 175, pp. 147–154, May 1983, doi: 10.1097/00003086-198305000-00021.

[59] D. A. Winter, *Biomechanics and Motor Control of Human Movement*, 4th ed. Hoboken, N.J: Wiley, 2009.

[60] J. M. Donelan, R. Kram, and A. D. Kuo, "Mechanical Work for Step-to-Step Transitions is a Major Determinant of the Metabolic Cost of Human Walking," *J. Exp. Biol.*, vol. 205, pp. 3717–3727, 2002, doi: 10.1242/jeb.205.23.3717.

[61] J. M. Donelan, R. Kram, and A. D. Kuo, "Simultaneous Positive and Negative External Mechanical Work in Human Walking," *J. Biomech.*, vol. 35, no. 1, pp. 117–124, Jan. 2002, doi: 10.1016/S0021-9290(01)00169-5.

[62] A. D. Kuo, J. M. Donelan, and A. Ruina, "Energetic Consequences of Walking Like an Inverted Pendulum: Step-to-Step Transitions," *Exerc. Sport Sci. Rev.*, vol. 33, no. 2, pp. 88–97, Apr. 2005, doi: 10.1097/00003677-200504000-00006.

[63] K. Sasaki, R. R. Neptune, and S. A. Kautz, "The Relationships Between Muscle, External, Internal and Joint Mechanica Work during Normal Walking," *J. Exp. Biol.*, vol. 212, no. 5, pp. 738–744, Mar. 2009, doi: 10.1242/jeb.023267.

[64] G. A. Pratt and M. M. Williamson, "Series Elastic Actuators," in *Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots*, Pittsburgh, PA, USA, 1995, vol. 1, pp. 399–406. doi: 10.1109/IROS.1995.525827.

[65] E. Krimsky and S. H. Collins, "Optimal Control of an Energy-Recycling Actuator for Mobile Robotics Applications," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, May 2020, pp. 3559–3565. doi: 10.1109/ICRA40945.2020.9196870.

[66] E. Bolivar, S. Rezazadeh, T. Summers, and R. D. Gregg, "Robust Optimal Design of Energy Efficient Series Elastic Actuators: Application to a Powered Prosthetic Ankle," in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, Toronto, ON, Canada, Jun. 2019, pp. 740–747. doi: 10.1109/ICORR.2019.8779446.

[67] E. A. Bolivar Nieto, S. Rezazadeh, and R. D. Gregg, "Minimizing Energy Consumption and Peak Power of Series Elastic Actuators: A Convex Optimization Framework for Elastic Element Design," *IEEEASME Trans. Mechatron.*, vol. 24, no. 3, pp. 1334–1345, Jun. 2019, doi: 10.1109/TMECH.2019.2906887.

[68] E. A. Bolıvar-Nieto, G. C. Thomas, E. Rouse, and R. D. Gregg, "Convex Optimization for Spring Design in Series Elastic Actuators: From Theory to Practice," in *International Conference on Intelligent Robots and Systems (IROS)*, Prague, Czech Republic, 2021, p. 6.

[69] E. Bolívar, S. Rezazadeh, and R. Gregg, "A General Framework for Minimizing Energy Consumption of Series Elastic Actuators With Regeneration," in *ASME 2017 Dynamic Systems and Control Conference*, Tysons, Virginia, USA, Oct. 2017, p. V001T36A005. doi: 10.1115/DSCC2017-5373.

[70] E. J. Rouse, L. M. Mooney, and H. M. Herr, "Clutchable Series-Elastic Actuator: Implications for Prosthetic Knee Design," *Int. J. Robot. Res.*, vol. 33, no. 13, pp. 1611–1625, Nov. 2014, doi: 10.1177/0278364914545673.

[71] E. C. Martinez-Villalpando and H. Herr, “Agonist-Antagonist Active Knee Prosthesis: A Preliminary Study in Level-Ground Walking,” *J. Rehabil. Res. Dev.*, vol. 46, no. 3, pp. 361–374, 2009, doi: 10.1682/JRRD.2008.09.0131.

[72] E. C. Martinez-Villalpando, L. Mooney, G. Elliott, and H. Herr, “Antagonistic Active Knee Prosthesis. A Metabolic Cost of Walking Comparison with a Variable-Damping Prosthetic Knee,” in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, MA, Aug. 2011, pp. 8519–8522. doi: 10.1109/IEMBS.2011.6092102.

[73] L. Mooney and H. Herr, “Continuously-Variable Series-Elastic Actuator,” in *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*, Seattle, WA, Jun. 2013, pp. 1–6. doi: 10.1109/ICORR.2013.6650402.

[74] S. Seok *et al.*, “Design Principles for Energy-Efficient Legged Locomotion and Implementation on the MIT Cheetah Robot,” *IEEEASME Trans. Mechatron.*, vol. 20, no. 3, pp. 1117–1129, Jun. 2015, doi: 10.1109/TMECH.2014.2339013.

[75] A. Wang and S. Kim, “Directional Efficiency in Geared Transmissions: Characterization of Backdrivability Towards Improved Proprioceptive Control,” in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, WA, USA, May 2015, pp. 1055–1062. doi: 10.1109/ICRA.2015.7139307.

[76] S. Seok, A. Wang, D. Otten, and S. Kim, “Actuator Design for High Force Proprioceptive Control in Fast Legged Locomotion,” in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2012)*, Vilamoura-Algarve, Portugal, Oct. 2012, pp. 1970–1975. doi: 10.1109/IROS.2012.6386252.

[77] S. Seok, A. Wang, Meng Yee Chuah, D. Otten, J. Lang, and S. Kim, “Design Principles for Highly Efficient Quadrupeds and Implementation on the MIT Cheetah Robot,” in *2013 IEEE International Conference on Robotics and Automation (ICRA)*, Karlsruhe, Germany, May 2013, pp. 3307–3312. doi: 10.1109/ICRA.2013.6631038.

[78] P. M. Wensing, A. Wang, S. Seok, D. Otten, J. Lang, and S. Kim, “Proprioceptive Actuator Design in the MIT Cheetah: Impact Mitigation and High-Bandwidth Physical Interaction for Dynamic Legged Robots,” *IEEE Trans. Robot.*, vol. 33, no. 3, pp. 509–522, Jun. 2017, doi: 10.1109/TRO.2016.2640183.

[79] H. Zhu, J. Doan, C. Stence, G. Lv, T. Elery, and R. Gregg, “Design and Validation of a Torque Dense, Highly Backdrivable Powered Knee-Ankle Orthosis,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, Singapore, May 2017, pp. 504–510. doi: 10.1109/ICRA.2017.7989063.

[80] H. Zhu, C. Nesler, N. Divekar, V. Peddinti, and R. Gregg, “Design Principles for Compact, Backdrivable Actuation in Partial-Assist Powered Knee Orthoses,” *IEEEASME Trans. Mechatron.*, pp. 1–1, 2021, doi: 10.1109/TMECH.2021.3053226.

[81] H. Zhu, C. Nesler, N. Divekar, M. T. Ahmad, and R. D. Gregg, “Design and Validation of a Partial-Assist Knee Orthosis with Compact, Backdrivable Actuation,” in *2019 IEEE 16th*

International Conference on Rehabilitation Robotics (ICORR), Toronto, ON, Canada, Jun. 2019, pp. 917–924. doi: 10.1109/ICORR.2019.8779479.

[82] G. Lv, H. Zhu, and R. D. Gregg, "On the Design and Control of Highly Backdrivable Lower-Limb Exoskeletons: A Discussion of Past and Ongoing Work," *IEEE Control Syst.*, vol. 38, no. 6, pp. 88–113, Dec. 2018, doi: 10.1109/MCS.2018.2866605.

[83] T. Elery, S. Rezazadeh, C. Nesler, and R. D. Gregg, "Design and Validation of a Powered Knee–Ankle Prosthesis With High-Torque, Low-Impedance Actuators," *IEEE Trans. Robot.*, vol. 36, no. 6, pp. 1649–1668, Dec. 2020, doi: 10.1109/TRO.2020.3005533.

[84] T. Elery, S. Rezazadeh, C. Nesler, J. Doan, H. Zhu, and R. D. Gregg, "Design and Benchtop Validation of a Powered Knee-Ankle Prosthesis with High-Torque, Low-Impedance Actuators," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, QLD, May 2018, pp. 2788–2795. doi: 10.1109/ICRA.2018.8461259.

[85] C. Nesler, G. Thomas, N. Divekar, E. J. Rouse, and R. D. Gregg, "Enhancing Voluntary Motion with Modular, Backdrivable, Powered Hip and Knee Orthoses," *ArXiv211001562 Cs*, Oct. 2021, Accessed: Oct. 10, 2021. [Online]. Available: http://arxiv.org/abs/2110.01562

[86] J. M. Donelan, Q. Li, V. Naing, J. A. Hoffer, D. J. Weber, and A. D. Kuo, "Biomechanical Energy Harvesting: Generating Electricity During Walking with Minimal User Effort," *Science*, vol. 319, no. 5864, pp. 807–810, Feb. 2008, doi: 10.1126/science.1149860.

[87] J. M. Donelan, V. Naing, and Q. Li, "Biomechanical Energy Harvesting," in *2009 IEEE Radio and Wireless Symposium*, San Diego, CA, USA, Jan. 2009, p. 4. doi: 10.1109/RWS.2009.4957269.

[88] Q. Li, V. Naing, J. A. Hoffer, D. J. Weber, A. D. Kuo, and J. M. Donelan, "Biomechanical Energy Harvesting: Apparatus and Method," in *2008 IEEE International Conference on Robotics and Automation (ICRA)*, Pasadena, CA, USA, May 2008, pp. 3672–3677. doi: 10.1109/ROBOT.2008.4543774.

[89] Q. Li, V. Naing, and J. M. Donelan, "Development of a Biomechanical Energy Harvester," *J. NeuroEngineering Rehabil.*, vol. 6, no. 1, p. 22, Dec. 2009, doi: 10.1186/1743-0003-6-22.

[90] J. C. Selinger and J. M. Donelan, "Myoelectric Control for Adaptable Biomechanical Energy Harvesting," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 3, pp. 364–373, Mar. 2016, doi: 10.1109/TNSRE.2015.2510546.

[91] T. Barto and D. Simon, "Neural Network Control of an Optimized Regenerative Motor Drive for a Lower-Limb Prosthesis," in *2017 American Control Conference (ACC)*, Seattle, WA, USA, May 2017, pp. 5330–5335. doi: 10.23919/ACC.2017.7963783.

[92] A. Ghorbanpour and H. Richter, "Control With Optimal Energy Regeneration in Robot Manipulators Driven by Brushless DC Motors," in *ASME 2018 Dynamic Systems and Control Conference*, Atlanta, Georgia, USA, Sep. 2018, p. V001T04A003. doi: 10.1115/DSCC2018-8972.

[93] A. Ghorbanpour and H. Richter, "Energy-Optimal, Direct-Phase Control of Brushless Motors for Robotic Drives," in *ASME 2020 Dynamic Systems and Control Conference*, Virtual, Online, Oct. 2020, p. V001T05A006. doi: 10.1115/DSCC2020-3200.

[94] R. Rarick, H. Richter, A. van den Bogert, D. Simon, H. Warner, and T. Barto, “Optimal Design of a Transfemoral Prosthesis with Energy Storage and Regeneration,” in *2014 American Control Conference (ACC)*, Portland, OR, USA, Jun. 2014, pp. 4108–4113. doi: 10.1109/ACC.2014.6859051.

[95] P. Khalaf and H. Richter, “Parametric Optimization of Stored Energy in Robots with Regenerative Drive Systems,” in *2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, Banff, AB, Canada, Jul. 2016, pp. 1424–1429. doi: 10.1109/AIM.2016.7576970.

[96] H. Richter, “A Framework for Control of Robots With Energy Regeneration,” *ASME J. Dyn. Syst. Meas. Control*, vol. 137, no. 9, p. 091004, Sep. 2015, doi: 10.1115/1.4030391.

[97] E. G. dos Santos and H. Richter, “Modeling and Control of a Novel Variable-Stiffness Regenerative Actuator,” in *ASME 2018 Dynamic Systems and Control Conference*, Atlanta, Georgia, USA, Sep. 2018, p. V002T24A003. doi: 10.1115/DSCC2018-9054.

[98] G. Khademi, H. Richter, and D. Simon, “Multi-Objective Optimization of Tracking/Impedance Control for a Prosthetic Leg with Energy Regeneration,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Las Vegas, NV, USA, Dec. 2016, pp. 5322–5327. doi: 10.1109/CDC.2016.7799085.

[99] H. Warner, D. Simon, and H. Richter, “Design Optimization and Control of a Crank-Slider Actuator for a Lower-Limb Prosthesis with Energy Regeneration,” in *2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, Banff, AB, Canada, Jul. 2016, pp. 1430–1435. doi: 10.1109/AIM.2016.7576971.

[100] B.-H. Kim and H. Richter, “Energy Regeneration-Based Hybrid Control for Transfemoral Prosthetic Legs Using Four-Bar Mechanism,” in *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, Washington, DC, USA, Oct. 2018, pp. 2516–2521. doi: 10.1109/IECON.2018.8591399.

[101] G. Khademi, H. Mohammadi, H. Richter, and D. Simon, “Optimal Mixed Tracking/Impedance Control With Application to Transfemoral Prostheses With Energy Regeneration,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 4, pp. 894–910, Apr. 2018, doi: 10.1109/TBME.2017.2725740.

[102] F. Rohani, H. Richter, and A. J. van den Bogert, “Optimal Design and Control of an Electromechanical Transfemoral Prosthesis with Energy Regeneration,” *PLOS ONE*, vol. 12, no. 11, p. e0188266, Nov. 2017, doi: 10.1371/journal.pone.0188266.

[103] P. Khalaf and H. Richter, “On Global, Closed-Form Solutions to Parametric Optimization Problems for Robots With Energy Regeneration,” *ASME J. Dyn. Syst. Meas. Control*, vol. 140, no. 3, p. 031003, Mar. 2018, doi: 10.1115/1.4037653.

[104] P. Khalaf, H. Warner, E. Hardin, H. Richter, and D. Simon, “Development and Experimental Validation of an Energy Regenerative Prosthetic Knee Controller and Prototype,” in *ASME 2018 Dynamic Systems and Control Conference*, Atlanta, Georgia, USA, Sep. 2018, p. V001T07A008. doi: 10.1115/DSCC2018-9091.

[105] A. Ghorbanpour and H. Richter, “A Novel Concept for Energy-Optimal, Independent-Phase Control of Brushless Motor Drivers,” *ASME Lett. Dyn. Syst. Control*, pp. 1–11, Oct. 2021, doi: 10.1115/1.4052662.

[106] J. Andrysek and G. Chau, “An Electromechanical Swing-Phase-Controlled Prosthetic Knee Joint for Conversion of Physiological Energy to Electrical Energy: Feasibility Study,” *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2276–2283, Dec. 2007, doi: 10.1109/TBME.2007.908309.

[107] J. Andrysek, T. Liang, and B. Steinnagel, “Evaluation of a Prosthetic Swing-Phase Controller With Electrical Power Generation,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 17, no. 4, pp. 390–396, Aug. 2009, doi: 10.1109/TNSRE.2009.2023292.

[108] M. R. Tucker and K. B. Fite, “Mechanical Damping with Electrical Regeneration for a Powered Transfemoral Prosthesis,” in *2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, Montreal, QC, Canada, Jul. 2010, pp. 13–18. doi: 10.1109/AIM.2010.5695828.

[109] R. Riemer, R. W. Nuckols, and G. S. Sawicki, “Extracting Electricity with Exosuit Braking,” *Science*, vol. 372, no. 6545, pp. 909–911, May 2021, doi: 10.1126/science.abh4007.

[110] E. Schertzer and R. Riemer, “Harvesting Biomechanical Energy or Carrying Batteries? An Evaluation Method based on a Comparison of Metabolic Power,” *J. NeuroEngineering Rehabil.*, vol. 12, no. 1, p. 30, Dec. 2015, doi: 10.1186/s12984-015-0023-7.

[111] R. Riemer and A. Shapiro, “Biomechanical Energy Harvesting from Human Motion: Theory, State of the Art, Design Guidelines, and Future Directions,” *J. NeuroEngineering Rehabil.*, vol. 8, no. 1, p. 22, 2011, doi: 10.1186/1743-0003-8-22.

[112] M. S. Orendurff, J. Schoen, G. Bernatz, A. Segal, and G. Klute, “How Humans Walk: Bout Duration, Steps per Bout, and Rest Duration,” *J. Rehabil. Res. Dev.*, vol. 45, no. 7, pp. 1077–1090, Dec. 2008, doi: 10.1682/JRRD.2007.11.0197.

[113] Y. Feng, J. Mai, S. K. Agrawal, and Q. Wang, “Energy Regeneration From Electromagnetic Induction by Human Dynamics for Lower Extremity Robotic Prostheses,” *IEEE Trans. Robot.*, vol. 36, no. 5, pp. 1442–1451, Oct. 2020, doi: 10.1109/TRO.2020.2991969.

[114] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in *Advances in Neural Information Processing Systems (NIPS)*, 2012, vol. 25, pp. 1097–1105.

[115] O. Russakovsky *et al.*, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.

[116] V. Rai and E. Rombokas, “Evaluation of a Visual Localization System for Environment Awareness in Assistive Devices,” in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, USA, Jul. 2018, pp. 5135–5141. doi: 10.1109/EMBC.2018.8513442.

[117] M. Abadi *et al.*, “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems,” *arXiv:1603.04467*, Mar. 2016, Accessed: Dec. 09, 2021. [Online]. Available: http://arxiv.org/abs/1603.04467

[118] X. Glorot and Y. Bengio, “Understanding the Difficulty of Training Deep Feedforward Neural Networks,” in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, Sardinia, Italy, 2010, vol. 9, pp. 249–256.

[119] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *arXiv*, Dec. 2014, doi: arXiv:1412.6980.

[120] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A Large-Scale Hierarchical Image Database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, USA, p. 8. doi: 10.1109/CVPR.2009.5206848.

[121] M. Li *et al.*, “Gaze Fixation Comparisons Between Amputees and Able-bodied Individuals in Approaching Stairs and Level-ground Transitions: A Pilot Study,” in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, Jul. 2019, pp. 3163–3166. doi: 10.1109/EMBC.2019.8857388.

[122] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” *arXiv*, May 2019, doi: arXiv:1905.11946.

[123] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” *arXiv*, Dec. 2015, doi: arXiv:1512.00567.

[124] A. G. Howard *et al.*, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” *arXiv*, Apr. 2017, doi: arXiv:1704.04861.

[125] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” *arXiv*, Jan. 2018, doi: arXiv:1801.04381.

[126] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *arXiv*, Sep. 2014, doi: arXiv:1409.1556.

[127] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” *arXiv*, Oct. 2016, doi: arXiv:1610.02357.

[128] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *arXiv*, Dec. 2015, doi: arXiv:1512.03385.

[129] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” *arXiv*, Aug. 2016, doi: arXiv:1608.06993.

[130] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.

[131] I. Loshchilov and F. Hutter, “SGDR: Stochastic Gradient Descent with Warm Restarts,” *arXiv*, Aug. 2016, doi: arXiv:1608.03983.

[132] X. Ding, X. Zhang, N. Ma, J. Han, G. Ding, and J. Sun, “RepVGG: Making VGG-style ConvNets Great Again,” *arXiv*, Jan. 2021, doi: arXiv:2101.03697.

[133] A. Canziani, A. Paszke, and E. Culurciello, “An Analysis of Deep Neural Network Models for Practical Applications,” *arXiv*, May 2016, doi: arXiv:1605.07678.

[134] A. Wong, “NetScore: Towards Universal Metrics for Large-scale Performance Analysis of Deep Neural Networks for Practical On-Device Edge Usage,” *arXiv*, Jun. 2018, doi: arXiv:1806.05512.

[135] B. Laschowski, W. McNally, A. Wong, and J. McPhee, “Computer Vision and Deep Learning for Environment-Adaptive Control of Robotic Lower-Limb Exoskeletons,” Apr. 2021. doi: 10.1101/2021.04.02.438126.

[136] B. Laschowski, W. McNally, A. Wong, and J. McPhee, “Computer Vision and Deep Learning for Environment-Adaptive Control of Robotic Lower-Limb Exoskeletons,” in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Mexico, Nov. 2021, pp. 4631–4635. doi: 10.1109/EMBC46164.2021.9630064.

[137] F. Gao, F. Zhang, and H. Huang, “Investigation of Sit-to-Stand and Stand-to-Sit in an Above Knee Amputee,” in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Boston, MA, USA, Aug. 2011, pp. 7340–7343. doi: 10.1109/IEMBS.2011.6091712.

[138] M. K. Shepherd and E. J. Rouse, “Design and Characterization of a Torque-Controllable Actuator for Knee Assistance during Sit-to-Stand,” in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, USA, Aug. 2016, pp. 2228–2231. doi: 10.1109/EMBC.2016.7591172.

[139] M. K. Shepherd and E. J. Rouse, “Design and Validation of a Torque-Controllable Knee Exoskeleton for Sit-to-Stand Assistance,” *IEEEASME Trans. Mechatron.*, vol. 22, no. 4, pp. 1695–1704, Aug. 2017, doi: 10.1109/TMECH.2017.2704521.

[140] A. M. Simon, N. P. Fey, K. A. Ingraham, S. B. Finucane, E. G. Halsne, and L. J. Hargrove, “Improved Weight-Bearing Symmetry for Transfemoral Amputees During Standing Up and Sitting Down With a Powered Knee-Ankle Prosthesis,” *Arch. Phys. Med. Rehabil.*, vol. 97, no. 7, pp. 1100–1106, Jul. 2016, doi: 10.1016/j.apmr.2015.11.006.

[141] J. Skelton, S.-K. Wu, and X. Shen, “Design of a Powered Lower-Extremity Orthosis for Sit-to-Stand and Ambulation Assistance,” *ASME J. Med. Devices*, vol. 7, no. 3, p. 030910, Sep. 2013, doi: 10.1115/1.4024489.

[142] J. Vantilt *et al.*, “Model-Based Control for Exoskeletons with Series Elastic Actuators Evaluated on Sit-to-Stand Movements,” *J. NeuroEngineering Rehabil.*, vol. 16, no. 1, p. 65, Jun. 2019, doi: 10.1186/s12984-019-0526-8.

[143] H. A. Varol, F. Sup, and M. Goldfarb, “Powered Sit-to-Stand and Assistive Stand-to-Sit Framework for a Powered Transfemoral Prosthesis,” in *2009 IEEE International Conference on Rehabilitation Robotics (ICORR)*, Kyoto, Japan, Jun. 2009, pp. 645–651. doi: 10.1109/ICORR.2009.5209582.

[144] M. Wu, M. R. Haque, and X. Shen, “Sit-to-Stand Control of Powered Knee Prostheses,” in *2017 ASME Design of Medical Devices Conference*, Minneapolis, Minnesota, USA, Apr. 2017, p. V001T05A015. doi: 10.1115/DMD2017-3507.

[145] D. Raz, E. Bolıvar-Nieto, N. Ozay, and R. D. Gregg, “Toward Phase-Variable Control of Sit-to-Stand Motion with a Powered Knee-Ankle Prosthesis,” in *IEEE Conference on Control Technology and Applications (CCTA)*, San Diego, CA, USA, 2021, p. 7.

[146] A. M. Simon, N. P. Fey, K. A. Ingraham, A. J. Young, and L. J. Hargrove, “Powered Prosthesis Control during Walking, Sitting, Standing, and Non-Weight Bearing Activities using Neural and Mechanical Inputs,” in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, San Diego, CA, USA, Nov. 2013, pp. 1174–1177. doi: 10.1109/NER.2013.6696148.

[147] S. Thapa, H. Zheng, G. F. Kogler, and X. Shen, “A Robotic Knee Orthosis for Sit-to-Stand Assistance,” in *ASME 2016 Dynamic Systems and Control Conference (DSCC)*, Minneapolis, Minnesota, USA, Oct. 2016, p. V001T07A004. doi: 10.1115/DSCC2016-9891.

[148] G. Wu *et al.*, “ISB Recommendation on Definitions of Joint Coordinate System of Various Joints for the Reporting of Human Joint Motion — Part I: Ankle, Hip, and Spine,” *J. Biomech.*, vol. 35, no. 4, pp. 543–548, Apr. 2002, doi: 10.1016/S0021-9290(01)00222-6.

[149] S. L. Delp *et al.*, “OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement,” *IEEE Trans. Biomed. Eng.*, vol. 54, no. 11, pp. 1940–1950, Nov. 2007, doi: 10.1109/TBME.2007.901024.

[150] A. J. van den Bogert and A. Su, “A Weighted Least Squares Method for Inverse Dynamic Analysis,” *Comput. Methods Biomech. Biomed. Engin.*, vol. 11, no. 1, pp. 3–9, Feb. 2008, doi: 10.1080/10255840701550865.

[151] B. Laschowski and J. McPhee, “Quantifying Body Segment Parameters Using Dual-Energy X-Ray Absorptiometry: A Paralympic Wheelchair Curler Case Report,” *Procedia Eng.*, vol. 147, pp. 163–167, 2016, doi: 10.1016/j.proeng.2016.06.207.

[152] B. Laschowski and J. McPhee, “Body Segment Parameters of Paralympic Athletes from Dual-Energy X-Ray Absorptiometry,” *Sports Eng.*, vol. 19, no. 3, pp. 155–162, Sep. 2016, doi: 10.1007/s12283-016-0200-3.

[153] R. J. Wong and J. Smith, “Regenerative Effects in the Sit-to-Stand and Stand-to-Sit Movement,” *Robotica*, vol. 33, no. 1, pp. 107–126, Jan. 2015, doi: 10.1017/S026357471400006X.

[154] M. Bortole *et al.*, “The H2 Robotic Exoskeleton for Gait Rehabilitation after Stroke: Early Findings from a Clinical Study,” *J. NeuroEngineering Rehabil.*, vol. 12, no. 1, p. 54, Dec. 2015, doi: 10.1186/s12984-015-0048-y.

[155] W. Herzog, “The Relation Between the Resultant Moments at a Joint and the Moments Measured by an Isokinetic Dynamometer,” *J. Biomech.*, vol. 21, no. 1, pp. 5–12, Jan. 1988, doi: 10.1016/0021-9290(88)90185-6.

[156] U. H. Lee, C.-W. Pan, and E. J. Rouse, “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)*, Macau, China, Nov. 2019, pp. 8018–8025. doi: 10.1109/IROS40897.2019.8967626.

[157] N. E. Krausz and L. J. Hargrove, “A Survey of Teleceptive Sensing for Wearable Assistive Robotic Devices,” *Sensors*, vol. 19, no. 23, p. 5238, Nov. 2019, doi: 10.3390/s19235238.

[158] K. Zhang, C. W. de Silva, and C. Fu, “Sensor Fusion for Predictive Control of Human-Prosthesis-Environment Dynamics in Assistive Walking: A Survey,” *arXiv*, Mar. 2019, doi: arXiv:1903.07674.

[159] K. A. Inkol and J. McPhee, “Assessing Control of Fixed-Support Balance Recovery in Wearable Lower-Limb Exoskeletons Using Multibody Dynamic Modelling,” in *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, New York City, NY, USA, Nov. 2020, pp. 54–60. doi: 10.1109/BioRob49111.2020.9224430.

[160] A. Vaswani *et al.*, “Attention Is All You Need,” *arXiv*, Jun. 2017, doi: arXiv:1706.03762.

[161] Y. Bengio, P. Simard, and P. Frasconi, “Learning Long-Term Dependencies with Gradient Descent is Difficult,” *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994, doi: 10.1109/72.279181.

[162] A. F. Burke, “Batteries and Ultracapacitors for Electric, Hybrid, and Fuel Cell Vehicles,” *Proc. IEEE*, vol. 95, no. 4, pp. 806–820, Apr. 2007, doi: 10.1109/JPROC.2007.892490.

[163] B. Hunter, “Design of a Self-Contained, Active, Regenerative Computer Controlled Above-Knee Prosthesis,” PhD Thesis, Massachusetts Institute of Technology, 1981.

[164] B. Seth, “Energy Regeneration and Its Application to Active Above-Knee Prostheses,” PhD Thesis, Massachusetts Institute of Technology, 1987.

[165] K. Tabor, “The Real-Time Digital Control of a Regenerative Above-Knee Prosthesis,” PhD Thesis, Massachusetts Institute of Technology, 1988.