Studying the Biomechanics of a Wheelchair Basketball Free Throw using Pose Estimation

by

Hisham Mohammad

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2025

© Hisham Mohammad 2025

Author's Declaration

I hereby declare that I am the sole author of this 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.

Abstract

Wheelchair basketball is a popular Paralympic sport where athletes with varying disabilities compete under a point-based classification system. Lower-class athletes (1.0–2.5), with higher levels of disability, often struggle to engage their trunk and core muscles, while higher-class athletes (3.0–4.5) have greater functional ability and utilize their trunk extensively. Coaches must consider these functional disparities when formulating strategies and designing individualized training regimens.

Consistent free-throw shooting is critical in wheelchair basketball, as it offers an uncontested scoring opportunity. Higher-class athletes, who incorporate trunk motion, rely less on their arms for force generation, resulting in distinct shooting mechanics. Given the biomechanical variability arising from these physical differences, understanding individual shooting techniques is vital for optimizing performance.

Motion capture technologies are widely employed to analyze and improve athletic movements. However, traditional systems, such as wearable sensors and marker-based motion tracking, are often costly, time-intensive, and restrictive to mobility. Markerless motion capture systems address these limitations using computer vision techniques like pose estimation. Convolutional neural networks (CNNs) trained on large human image datasets can accurately detect joints and limbs, enabling real-time analysis. Commercial systems typically require multiple cameras, but deploying pose estimation CNNs on mobile devices allows motion analysis using only a built-in camera, enhancing portability and accessibility for sports training and biomechanical research.

This research focuses on designing and deploying pose estimation models within a mobile application to analyze the shooting arm's motion during a basketball free throw, with specific considerations for wheelchair basketball players. The pose estimation models, trained on a *COCO-WholeBody*^{TM} dataset to detect fingertip positions, were deployed on an iPhone and tested for accuracy and computational performance, particularly real-time motion analysis. The derived joint positions are used to calculate kinematic and dynamic metrics, including joint angles and torques.

The system's joint angle calculations were compared against the Vicon motion capture system. While upper arm and elbow angle errors had a root mean squared error (RMSE) within an acceptable range (less than 20°), wrist angle errors exceeded 65° due to limitations in pose estimation accuracy and the iPhone camera's frame rate. To demonstrate the system's utility, two shooting studies were conducted: (1) a comparison of biomechanics between one-motion and two-motion shooting techniques and (2) a biomechanical analysis of the shooting arm contrasting a national-level class 1 wheelchair basketball athlete with class 4.5 able-bodied participants shooting from a basketball wheelchair.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Prof. John McPhee for the continuous support of my MASc study and research, for his patience, motivation, guidance, and immense knowledge. His support during my internship allowed me to cultivate valuable engineering skills. I could not have imagined having a better advisor and mentor during my masters.

Next, I would like to thank Prof. Alex Wong for inspiring me to dive further into computer vision research and Prof. Steve Fischer for his ideas in motion capture analysis. Thanks to the current members and alumni of the Motion Research Group for their support during data collection.

This project would not have been possible without my friend, Amanda Hope, and her diligent work in app development. My research would not have come to life without your tireless efforts in troubleshooting errors in Swift.

I would also like to thank my former national wheelchair basketball coach, Michael Frogley. Your valuable insights and discussions greatly shaped the progress of my research and inspires me to continue my pursuit in developing impactful technology to enrich sports.

Finally, I would like to acknowledge the organizations and programs who supported this research including the University of Waterloo and the Canada Research Chairs (CRC) Program, and OnPara for providing a sport wheelchair.

Dedication

I would not be the person that I am without my parents, my sister, Dr. Faizah Ilyas. Thank you for your unwavering support and guidance.

Table of Contents

List of Figures ix

List of Tables xi

1 Introduction 1

1.1 Wheelchair Basketball . . . 1

1.1.1 Classification System . . . 2

1.1.2 Shooting Form based on Level of Disability . . . 3

1.2 Motion Capture Technologies in Sports . . . 5

1.3 Human Pose Estimation: Using Deep Learning for Markerless Motion Capture 6

1.3.1 Architecture of a CNN . . . 7

1.3.2 CNNs in Human Pose Estimation . . . 7

1.4 Research Objective . . . 8

1.5 Organization of Thesis . . . 9

2 Literature Review 10

2.1 Biomechanics of the Free Throw . . . 10

2.1.1 Free Throw Model Optimization . . . 11

2.1.2 Biomechanics of the Shooting Arm . . . 13

2.1.3 Free Throws in Wheelchair Basketball . . . 17

2.2 2D Human Pose Estimation . . . 18

2.2.1 Deep Learning Approaches . . . 19
2.2.2 Datasets and Evaluations . . . 23

3 Biomechanical Model of Basketball Shooting 26
3.1 Rationale for 2D Analysis . . . 26
3.2 Design of Model . . . 28
3.3 Kinematic Analysis . . . 31
3.4 Inverse Dynamics . . . 31
3.5 Validation . . . 33
3.6 Discussion . . . 34

4 Pose Estimation for Basketball Shooting 37
4.1 Computer Vision using Edge Computing . . . 37
4.1.1 Deployment of Computer Vision using Core ML . . . 37
4.1.2 MMPose Pose Estimation Models . . . 38
4.2 Implementation on an iPhone . . . 42
4.2.1 Walkthrough of *ShotPose* . . . 42
4.2.2 Data Processing . . . 44
4.3 Experiments with *ShotPose* . . . 44
4.3.1 Method . . . 45
4.3.2 Comparison to Marker-based System . . . 49
4.3.3 Assessing Wheelchair Basketball Free Throws using Motion Capture 54

5 Discussion 62
5.1 Comparison with State-of-the-Art Sports Pose Estimation Systems . . . 64
5.2 Future Research . . . 65
5.2.1 Performance of Pose Estimation . . . 65
5.2.2 Functionality of Free Throw Mobile App . . . 66
5.2.3 Interdisciplinary Collaboration . . . 68
5.2.4 Broader Vision . . . 69

References **70**

List of Figures

1.1 Graphical representation of the region of the body affected by a disability and the corresponding IWBF classification [87] 2

1.2 Low-class athletes lean against the backrest to compensate for lack of trunk stability [89] 3

1.3 High-class athletes can engage core muscles to assist with ball handling [89] 3

2.1 Parameters of the free throw [202] 12

2.2 Forces acting on a basketball after release [16] 12

2.3 Phases of the free throw and the corresponding joint movements [14] 14

2.4 One-motion shot 15

2.5 Two-motion shot 15

2.6 Set of arm joint positions at release [153] 16

2.7 Challenges of human pose estimation in images or videos. Examples are from the Max Planck Institute for Informatics (MPII) dataset [10] 19

2.8 (a) Regression method maps body joint coordinates from the input image. (b) Detection method predicts joint coordinates from representative heatmaps [220]. 20

2.9 Architecture of Simple Baselines: ResNet backbone with deconvolution layers for heatmap generation [205]. 21

2.10 Architecture of HRNet [181] 22

2.11 2D single person pose estimation datasets and metrics [220] 24

3.1 Geometric representation of the shooting arm during a free throw 29

3.2 Free body diagram of arm segments 32
3.3 Joint torque validation of Newton-Euler approach against MapleSim for the shoulder, elbow and wrist. Vertical line separates the transition phase from the launch phase. 36

4.1 Overview of the pose estimation algorithm 38
4.2 Wireframe of *ShotPose* in the IPhone XS template 43
4.3 Frame of a video before pose estimation 44
4.4 Pose estimation result with joint torques ($N \cdot m$) and limb segment angles (deg) 44
4.5 Red (solid): Vicon cameras, Yellow/black (dashed): iPhone XS 46
4.6 Left: Placement of the optical markers on the shooting arm. Right: Visualization of the shooting arm and position of the Vicon cameras in the Nexus software. 47
4.7 Definition of joint angles 51
4.8 Motion blur appearing in a frame of a free throw video 52
4.9 Wrist angle discrepancy between optical marker placement (green solid line) and detection of finger (red solid line) by the pose estimation model from *ShotPose* 53
4.10 Mean joint angles: Vicon vs. Free throw app. (Shaded region represents ± 1 std dev) 56
4.11 One-motion shot (left column) vs. Two-motion shot (right column). Vertical line (start of elbow extension) separates the transition phase from the launch phase. (Shaded region represents ± 1 std dev) 58
4.12 Differences between the class 1.0 (left column) and the class 4.5 shot (right column) shooting forms using pose estimation. (Shaded region represents ± 1 std dev) 60

List of Tables

1.1 Shooting function based on IWBF classifications [44] 4

3.1 Anthropometric parameters for torque calculations based on a male (M; 75.0 kg, 1.741 m) [48] 33

4.1 Performance metrics and sizes of MMPose models on the COCO-Wholebody dataset 41

4.2 Joint angle RMSE of *ShotPose* vs. Vicon 51

Chapter 1

Introduction

Wheelchair basketball is one of the most popular sports in the Paralympic Games. Sir Ludwig Guttmann, the driving force behind the Paralympic movement, was tasked with organizing a spinal cord injury center at Stoke Mandeville Hospital in England. At that time, para-sports served a rehabilitative purpose, enabling individuals with existing or newly acquired disabilities to develop strength and regain their independence. In 1948, Guttmann organized the first Stoke Mandeville Wheelchair Games. As several countries began participating in the Stoke Mandeville Games, its annual international presence grew, eventually becoming the predecessor to the modern-day Paralympic Games. Wheelchair basketball was one of eight sports that debuted at the first Paralympic Games in Rome, Italy, in 1960. In Tel Aviv, Israel, at the 1968 Paralympics, the women's wheelchair basketball competition was introduced [24].

1.1 Wheelchair Basketball

Wheelchair basketball shares many rules with able-bodied basketball. It is a 5-on-5 game where the objective is to score more points in the opponents' basket while preventing the opponents from scoring. Each athlete must use a sport wheelchair. Similar to able-bodied basketball, each net is 10 feet in height, each team has 8 seconds to cross the half-court line when inbounding from the backcourt, each offensive possession expires in 24 seconds, and an offensive player can remain in the key (the zone under the basket) for no longer than 3 seconds. Players may commit offensive and defensive fouls, which may grant the opposing team free throws, an uncontested shooting opportunity from the free throw line across from the net.

According to the International Wheelchair Basketball Federation (IWBF), athletes competing at an international level must have a diagnosed physical disability. In North America, however, able-bodied athletes can also play wheelchair basketball alongside their disabled teammates, but only up to a national level. Each athlete must be fully harnessed into a regulation basketball wheelchair. Basketball wheelchairs are identified by their characteristic cambered wheels and a horizontal crash bar around the feet. Each sports wheelchair is custom-fitted to the athlete's body and tailored to support their unique disability.

1.1.1 Classification System

Wheelchair basketball allows athletes with a range of disabilities to participate through a points-based classification system. An athlete is awarded a point, from 1.0 to 4.5, in 0.5 increments, based on their level of ability [20]. A class 1.0 athlete has the highest level of disability, with paralysis of the trunk/abdominal muscles from the mid-chest level and down, or significant upper-limb impairment. A class 4.5 athlete has the most physical ability, and are primarily trans-tibial or trans-femoral amputees. They have full functionality of their core, hip and leg muscles.

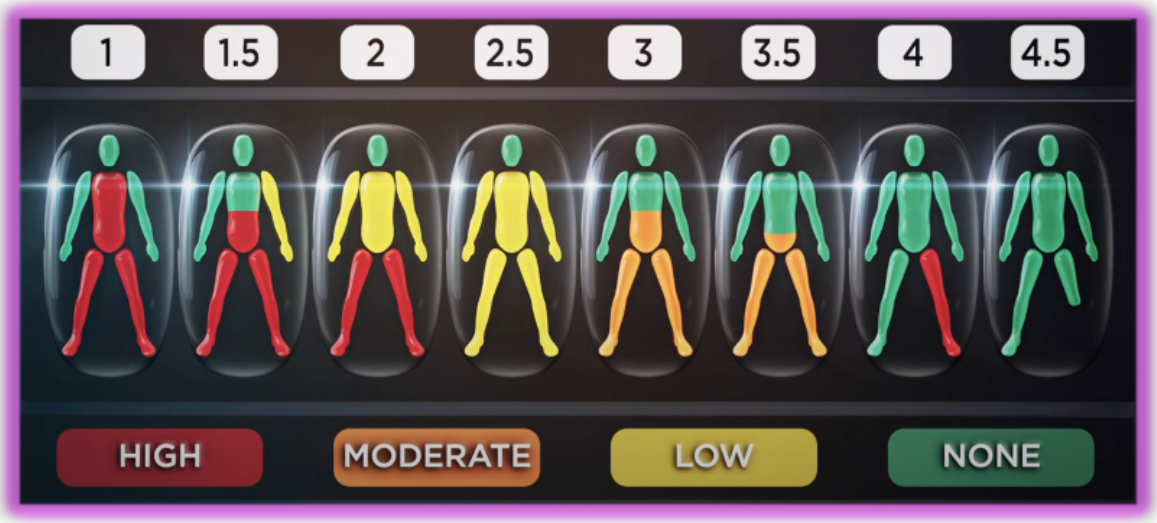


Figure 1.1: Graphical representation of the region of the body affected by a disability and the corresponding IWBF classification [87]

Low class athletes (class 1 or 2) generally sit lower to the ground with their knees higher than their hips and their backrests in a reclined position. This seating configuration provides a lower centre of gravity and compensates the athletes for the lack of trunk stability [215]. A higher class athlete's hips are level or noticeably higher than their knees. This allows the athlete to utilize their core muscles and legs when maneuvering the wheelchair. Higher class athletes are the more dominant players on either team. At any given time on the court, the sum of the classification points of the five athletes of each team cannot exceed 14 points. This allows athletes with different abilities to have the same opportunity to compete. It is the responsibility of the coach to strategically select the athletes based on their classification and disability.

1.1.2 Shooting Form based on Level of Disability

The level of disability of athletes affects the balance of their upper-body when handling the ball. It is easier for higher class athletes to pass and shoot a basketball because they can engage their abdominal and back muscles to sit upright. Lower class athletes must rely on their arms more when passing or shooting due to limitations in their trunk muscles (Table 1.1).



Figure 1.2: Low-class athletes lean against the backrest to compensate for lack of trunk stability [89]



Figure 1.3: High-class athletes can engage core muscles to assist with ball handling [89]

Professional basketball athletes refine their shooting form by adjusting the launch speed and angle to optimize each shot based on the height of release and other variables [193].

Class 1 - 1.5	Class 2 - 2.5	Class 3 - 4.5
Significant loss of stability in the trunk as the shooting arm is extended over the head during follow-through, often requiring arm support following the shot. During a two-handed shot, the trunk makes contact with the back of the wheelchair.	Mild to moderate loss of stability in the lower trunk during arm elevation and follow through, resulting in movement of the lower trunk away from the back of the wheelchair.	Excellent stability of the trunk while sitting upright, particularly in follow-through of the shot. The trunk moves toward the basket with shooting movement, without loss of stability.

Table 1.1: Shooting function based on IWBF classifications [44]

Able-bodied basketball players develop unique shooting styles depending on their ability to control key release parameters [16]. Some players may prefer a higher throwing angle to allow for greater margin of error in velocity, while others, with better velocity control, may opt for a flatter trajectory.

For example, a typical 6'3" player shoots at an average angle of 54.6 degrees from horizontal. However, Stephen Curry, one of the best free-throw shooters of this height, shoots at a steeper angle of 58.1 degrees. In contrast, James Harden, who is 6'5", shoots at a flatter angle of 49.6 degrees, compared to the NBA average of 53.4 degrees for players of his height [16].

Variations in wheelchair configuration and the physical abilities of the athlete contribute to the uniqueness of shooting forms in wheelchair basketball. Coaches often use the free throw as a benchmark for evaluating an athlete's shooting mechanics [82]. As the only uncontested scoring opportunity in basketball, the free throw provides a consistent scenario for developing proper shooting technique and can be easily replicated in practice.

Traditionally, wheelchair basketball coaches rely on empirical methods to analyze shooting form, focusing on the arc of the basketball and the consistency of the follow-through. However, these methods offer limited insight into the coordination of the shooting arm during the shot and do not provide quantitative data. To gain a more objective understanding of the athlete's motion, coaches are increasingly adopting motion capture technology.

1.2 Motion Capture Technologies in Sports

Modern motion capture technology can be broadly categorized into optical and non-optical systems. Non-optical systems, such as inertial measurement units (IMUs), rely on sensors attached to specific body parts. IMUs (*e.g.*, XSens MVN) measure the triaxial acceleration and angular velocity of body segments using accelerometers and gyroscopes, respectively [144].

Optical motion capture systems (*e.g.*, OptiTrack, Qualisys, Vicon) determine the 3D position of reflective markers placed on the subject's body by triangulating data captured by multiple infrared cameras. Marker-based optical systems are widely regarded as the "gold standard" for motion capture; however, they have notable limitations.

Limitations of IMUs and Marker-based Systems

Although highly portable, IMU measurements are prone to errors. Even when stationary, accelerometers and gyroscopes produce nonzero outputs, resulting in offset errors. Data drift is another common issue caused by the continuous integration of input signals to determine angular rotation from angular velocity. The input signal contains two primary components that contribute to drift: DC offset and noise. Integration of a nonzero constant from the DC offset leads to errors that grow linearly over time, while integration of noise, even with a zero mean, results in random walk [120]. Additionally, electromagnetic disturbances from the environment can interfere with measurement signals.

Marker-based motion capture systems, while considered the "gold standard," are less portable, restrict the subject's movement, and are prohibitively expensive. Conducting data capture sessions with these systems outside a controlled laboratory environment is challenging. In sports biomechanics, athletes may alter their natural movements due to restricted movement spaces or the awareness of markers on their body.

The process of placing markers on a participant's body is time-consuming and prone to placement errors. Such errors arise from incorrect placement of markers on anatomical landmarks or inconsistent placement between researchers [135]. These inaccuracies skew the orientation of segment anatomical coordinate systems and joint axes, leading to errors in joint angle calculations. For example, inconsistent marker placement has been shown to contribute up to 5 degrees of error in lower limb joint angles [74, 174].

Marker placement errors also stem from the displacement and deformation of clothing or soft tissue relative to the underlying skeletal anatomy during muscle contraction and movement [113]. When markers are placed on tight-fitting clothing or directly on the skin

(ideally), soft tissue artifacts conflict with the human body model's assumption of segment rigidity. This causes body segment kinematics to be inaccurately represented by marker movement rather than the actual movement of the segment. These position errors persist even when using motion capture suits. Overall, soft tissue artifacts are regarded as the primary source of errors in marker-based systems [113].

Finally, marker-based systems are extremely expensive, requiring numerous cameras to ensure marker visibility. Occlusion can prevent accurate marker detection, and capturing larger spaces demands additional cameras. Each Vicon or Qualisys camera costs thousands of dollars, significantly increasing the system's overall expense.

1.3 Human Pose Estimation: Using Deep Learning for Markerless Motion Capture

Markerless motion capture aims to overcome the limitations of marker-based systems by leveraging computer vision techniques. Traditional human modeling approaches relied on hand-engineered features (*e.g.*, Histogram of Oriented Gradients (HOG) [45], Scale-Invariant Feature Transforms (SIFT) [123]) to extract local features from images for detecting body parts and joints [47, 60]. These methods employed graphical tree structures to approximate the configuration of limbs between connecting joints, a process known as pose estimation, which involves determining the spatial arrangement of a person's body in an image or video [58].

While hand-designed features were robust to variations in skin pigmentation, lighting, and clothing, they struggled to generalize across different body parts. Additionally, graph-based human models faced challenges in detecting complex poses due to the lack of spatial inference between body parts located far apart.

Advances in computer hardware technology, the proliferation of publicly available image datasets and the advent of *deep learning* has brought rise to implementation of *convolutional neural networks* (CNNs) in computer vision applications [109]. In brief, CNNs are more effective than classical image processing approaches in analyzing the overall spatial arrangement and posture of a person's body while simultaneously identifying localized relationships between adjacent joints and specific body parts. The performance of CNNs was further enhanced through the use of graphical processing units (GPUs) for large-scale training [110], which ultimately led to AlexNet winning the 2012 ImageNet competition [51], a benchmark challenge designed to advance image recognition by evaluating algorithms on their ability to classify and detect objects in a large-scale image database

1.3.1 Architecture of a CNN

CNNs have been instrumental in the rapid advancement and integration of human pose estimation in markerless motion capture. These networks are a type of artificial neural network composed of multiple layers that transform the input, whether it be a color (RGB) or monochrome image. The hallmark of a CNN is the convolutional layer, which contains a set of filters (or kernels) with parameters (weights) learned during training. These filters are applied to an image using discrete 2D convolutions [109] to produce feature maps. The weights determine how the image is filtered (e.g., blurring, sharpening, or edge detection).

The initial convolutional layers of a CNN detect low-level features in groups of pixels, such as edges or curves. As the data passes through successive layers, the network identifies higher-level features, such as larger shapes or spatial relationships between image components. CNNs employ a data-driven approach to learn the optimal kernel weights by minimizing a loss function during training. This optimization is achieved iteratively using backpropagation [155] and gradient descent algorithms.

A typical CNN consists of thousands of convolutional layers with billions of trainable parameters [206]. Early architectures, such as LeNet [115] and VGG-Net [176], featured sequential stacks of convolutional layers. However, recent advances in computer vision research have led to significantly more complex CNN architectures, incorporating parallel branches of convolutions, multi-scale structures, and forward/skip connections. As a result, these neural networks are often limited to environments with high computational capabilities, including sufficient memory, bandwidth, and processing power.

To address these challenges, current research in deep learning focuses on designing more computationally efficient CNN architectures. Techniques such as quantization [221] and pruning [142] are commonly employed to reduce model size and complexity without significantly compromising performance.

1.3.2 CNNs in Human Pose Estimation

CNNs are widely used in human pose estimation to identify the location of joints and other anatomical landmarks, referred to as keypoints. Traditionally, the localization of keypoints has been performed to estimate poses in 2D.

Deep learning-based single-person pose estimation methods can be broadly categorized into *regression-based methods* and *detection-based methods*. Regression-based methods employ an end-to-end framework to directly map the input image to body joint coordinates or the parameters of human body models [191]. In contrast, the more commonly used

detection-based methods predict the locations of body parts [33] or joints [149] by representing them as rectangular image patches [90] or heatmaps. These heatmaps typically represent probable joint locations using a 2D Gaussian distribution [149], with the peak value indicating the most likely joint (keypoint) location. The objective of this approach is to predict a predefined number of heatmaps and minimize the error between the annotated ground truth and the peak of the predicted heatmap.

Multi-person pose estimation presents significantly greater challenges than single-person pose estimation due to its two-stage nature. It can be classified into *top-down* and *bottom-up* approaches. Top-down methods begin by using an off-the-shelf person detector to localize individuals within bounding boxes and then apply single-person pose estimation techniques to each box. In contrast, bottom-up methods detect all joint keypoints across the entire image and subsequently group them by person. Bottom-up methods are generally faster than top-down approaches because their inference time does not depend on the number of people in the input image.

1.4 Research Objective

Wheelchair basketball is a challenging sport to coach due to the complexity in the level of abilities across the athletes. An elite coach must be receptive to the athlete's disability when providing feedback to improve performance [54]. Pertaining to shooting, one of the most fundamental skills in basketball, a coach can only subjectively assess the technique without the use of technology. An athlete's shooting success can be better improved through diligent training using a scientific approach [19].

Motion capture technology can help in quantifying the motion of the shooting arm but has significant drawbacks, mainly due to expense and invasiveness. Markerless alternatives do not hamper the movement of a participant, but current systems like Theia3D consist of multiple expensive cameras, which prohibit in-the-wild application.

A 2D sagittal perspective is typically used to analyze the mechanics of a free throw. For a 2D application, it is more practical to use a monocular motion capture approach to avoid redundancy presented by a multi-camera system.

While 3D analysis provides comprehensive spatial information, a 2D approach was selected for this research based on several practical and theoretical considerations that align with the specific requirements of wheelchair basketball free throw analysis and the intended application environment.

This research proposes a system that performs 2D human pose estimation on the shooting arm of a wheelchair basketball athlete during a free throw. It allows coaches to analyze the kinematics and dynamics of the arm and correct shooting mechanics more objectively. Additionally, the pose estimation CNN employed by the motion capture system can be deployed to a mobile device with an integrated video camera to improve usability. This allows a coach to seamlessly augment athlete training through a convenient mobile application on their smartphone.

1.5 Organization of Thesis

To the best of the author's knowledge, the deployment of a mobile-based markerless motion capture system for biomechanics analysis of the basketball free throw has not been previously published. This thesis would pioneer a new form of coaching and the development of training regimes in wheelchair basketball and other para-sports. Furthermore, the pose-estimation model in the proposed system could be adapted for use in other sports and para-sports to analyze and improve biomechanical techniques, such as sprinting/wheelchair racing, swimming, or throwing, thereby enhancing performance and reducing injury risk. The outline of this thesis is as follows:

In Chapter 2, the relevant background for this thesis is provided, which include the biomechanics of a basketball free throw, and 2D human pose estimation.

In Chapter 3, the biomechanical models used to analyze the shooting arm of the free throw are discussed.

In Chapter 4, the details of CNN deployment onto a mobile device, the framework of the markerless motion capture smartphone application, and the results from the pose estimation experiments for wheelchair basketball free throw shooting are described.

Finally, chapter 5 concludes the thesis and documents future directions for this research.

Chapter 2

Literature Review

2.1 Biomechanics of the Free Throw

The free-throw is the most unique method of attaining points in basketball. It is the only shot in basketball that presents players with a completely uncontested opportunity where they shoot from the same distance from the rim and their outcome is determined solely by their own skill. A player is awarded a free-throw upon being fouled during the act of shooting, when the opposing team commits a technical or flagrant foul, or when the defending team fouls in a penalty situation. Free throws are worth one point. Players with high free-throw conversion percentages are an asset to a team during close games. As a result, athletes spend countless hours refining their free-throw shooting mechanics and establishing consistency.

Literature mostly focuses on the characteristics of the free throw shot in able-bodied basketball with wheelchair basketball receiving little attention. The skills of shooting, passing, and dribbling are agnostic to both sports. Skill can developed through diligent training and practice, but is also heavily influenced by an athlete's physical attributes and ability [188]. Wheelchair basketball players are positioned at a lower height than able-bodied players, and would thus generate propulsive forces from the arms and upper body when shooting. It is essential to understand each athlete's unique mechanics of the movement to improve shooting consistency, especially involving the difference in their classification.

Organization of Sections

In essence, the basketball free throw is a classical mechanical optimization problem. The success of a shot is based on fine-tuning the speed, angle and backspin imparted on the basketball [19]. These release conditions are equally fundamental in wheelchair basketball and able-bodied basketball.

This review will primarily focus on three areas of the existing literature: modeling studies that have attempted to define the "optimal" release conditions and trajectory of a free throw (Section 2.1.1), the kinetic and kinematic variables of the human body that influence shot outcome (Section 2.1.2), and the characteristics of the free throw in wheelchair basketball (Section 2.1.3).

2.1.1 Free Throw Model Optimization

Variables of the Free Throw

Recent literature on free throw modeling has identified four main parameters that influence the outcome of a free throw: release height, launch angle, release speed, and backspin. The release height depends on a player's anthropometric data (height, arm, and hand length) as well as their shooting form. The launch angle is the angle between the horizontal plane (parallel to the ground) and the ball's trajectory at release. The release speed is calculated by dividing the displacement of the ball's center by a short time interval. Backspin refers to the rotational velocity applied to the ball by the shooting hand. Athletes work to control these parameters when fine-tuning their shooting technique.

The height of an NBA or IWBF basketball net is 10 feet. The basketball rim is 18 inches in diameter and extends 6 inches from the backboard via a metal connector. The free throw line is located 15 feet from the front of the backboard. The ideal free throw is a swish (where the ball passes through the net without touching the rim). A regulation men's basketball is 29.5 inches in circumference and must weigh between 20 and 22 ounces (the same for both able-bodied and wheelchair basketball). A visualization of the free throw parameters is presented in Figure 2.1.

There are four main forces that act on the ball during flight (Figure 2.2). (1) The force of gravity pulls the ball toward the ground. (2) Buoyancy counteracts gravity by about 1.5% [16]. (3) The drag force, caused by air resistance, acts in the opposite direction of the ball's velocity and depends on the ball's velocity, its cross-sectional surface area, and the air's density. (4) Magnus force, which opposes gravity, arises from the difference in air

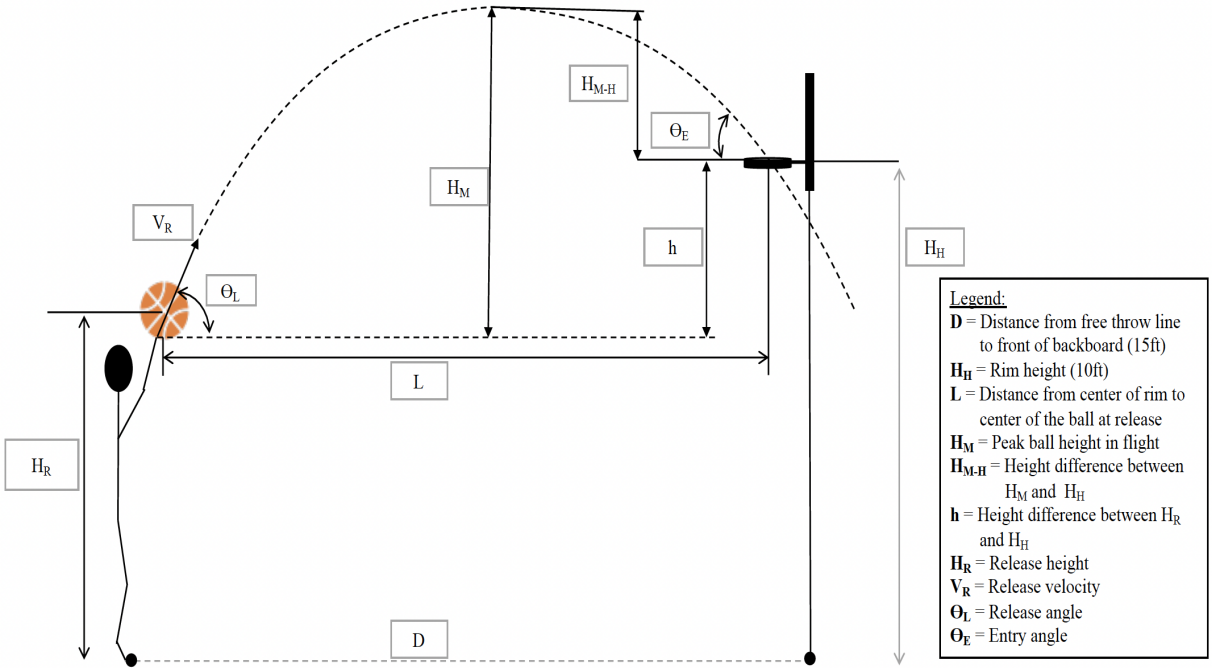


Figure 2.1: Parameters of the free throw [202]

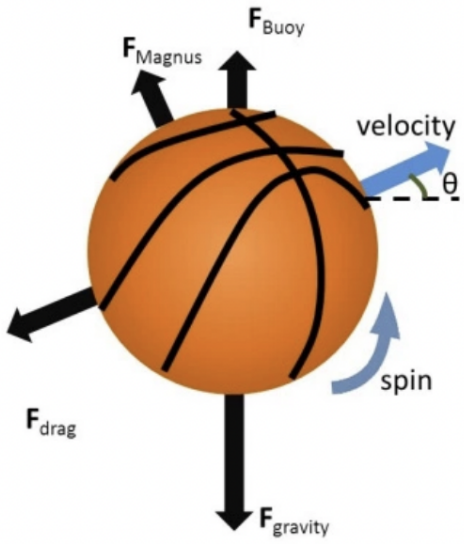


Figure 2.2: Forces acting on a basketball after release [16]

pressure above and below the ball due to backspin. According to [16], the ball's trajectory can be adequately represented by considering only buoyancy and gravity.

Models of the Free Throw

Early research by Brancazio [19] analyzed basketball shots using Newtonian mechanics, demonstrating that optimizing the shooting angle minimizes launch force, enabling efficient arm movement and reduced ball speed. Key findings included the sensitivity of shots to release velocity and angle, with optimal angles around $49 \pm 3°$.

Hamilton and Reinschmidt [75] expanded this work using computer simulations to optimize trajectories considering ball-rim and ball-backboard dynamics. They identified an optimal release angle of 59° and backspin of 1.5 rev/s from a 1 m release height, promoting the "back-rim-and-down" (BRAD) approach for enhanced error margins. While comprehensive, their sagittal plane model omitted lateral ball deviations, and a 3D analysis was recommended. The choice between 2D and 3D analysis approaches involves trade-offs between comprehensiveness and practical implementation, as will be discussed in Chapter 3.

Silverberg et al. [193] introduced a 3D model accounting for shooter skill levels. Optimal release conditions for swishes (angles 45°-55°, speed 6.8–7.16 m/s) and bank shots (same angles, speed 7.3–7.46 m/s) reinforced the BRAD approach. They emphasized a smooth motion and high release height for success.

Gablonsky and Lang [64] highlighted the dominance of release velocity over angle consistency in determining success. They showed that taller players benefit from larger margins of error and lower energy demands, with optimal angles decreasing from 57° to 52° as height increases.

Barzykina [16] incorporated Monte Carlo simulations and real shooting data for a predictive model tailored to individual biomechanics. This approach confirmed prior findings while accounting for variability in release conditions, enhancing model personalization.

2.1.2 Biomechanics of the Shooting Arm

Phases of the Free Throw

Motivated by Ozakazi et al.'s [152] effort in analyzing the jump shot, the free throw can be divided into three distinct phases. Since there is no jump or inertial follow-through in

a typical free throw, the inertial and stability phases from Ozakazi et al.'s study are not included. Figure 2.3 illustrates the preparatory (loading), transition, and launch phases. These phases represent the general movement patterns of a typical free throw, but each shooter will have their own unique form that may deviate from these phases. Wheelchair basketball athletes follow the same progression of phases during the free throw.

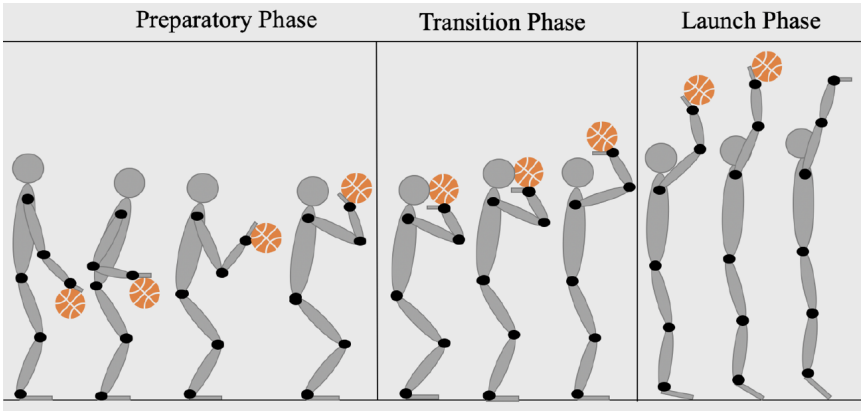


Figure 2.3: Phases of the free throw and the corresponding joint movements [14]

Preparatory Phase:
The shooter performs a unique pre-shot routine (dribbling the ball, adjusting stance, or aligning the wheelchair) after receiving the ball from the referee. Then, the player brings the ball to the waist using both hands, with their fingers spread and the fingertips making contact with the ball. In wheelchair basketball, the athlete either places the ball on their lap or starts to raise it on the same shoulder as their preferred shooting arm. The hand of the shooting arm supports the ball from underneath and generates the force that launches the ball, while the other hand acts as support. In able-bodied basketball, the feet are shoulder-width apart, whereas in wheelchair basketball, the chair is turned slightly to align the shooting shoulder with the net, with the casters just in front of the free throw line. In able-bodied basketball, the shooter lowers their center of mass through hip, knee, and ankle flexion [152]. In wheelchair basketball, athletes with trunk mobility start to extend their torso, while those with trunk muscle impairment recline against their backrest. Flexion of the shoulder and elbow occurs, and the wrist extends to stabilize the ball for propulsion. The end of the preparatory phase is marked by the stopping of the shoulder in the sagittal plane.

Transition Phase:
The transition phase is characterized by the elevation of the upper arm. The wrist tends to reach peak extension at the beginning of the transition phase. In able-bodied basketball,

the shooter extends their lower body in concert with shoulder extension. In wheelchair basketball, athletes with trunk functionality tend to lean forward during this phase. The end of the transition phase occurs with the onset of elbow extension and wrist flexion.

Launch Phase:
During the launch phase, elbow extension occurs in concert with rapid wrist flexion [152]. The coordination of the wrist and elbow is considered to be the most important factor in a successful shot [54]. Finger flexion provides backspin to the ball just before it leaves the hand. Shoulder flexion continues during this phase. In able-bodied basketball, the hip is fully extended and the ankle is plantarflexed for additional force [152]. Wheelchair basketball athletes with trunk functionality fully extend their trunk.

Types of Shots:
There are two prominent styles of shooting in basketball: one-motion and two-motion shots. Athletes adopting the one-motion shot do not stop the motion of the ball at the shoulder between the transition phase and the launch phase. Those who use the two-motion form will stop the ball at the shoulder before launching it (see Figures 2.4 and 2.5).

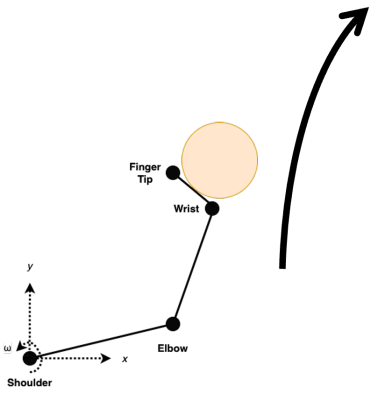


Figure 2.4: One-motion shot

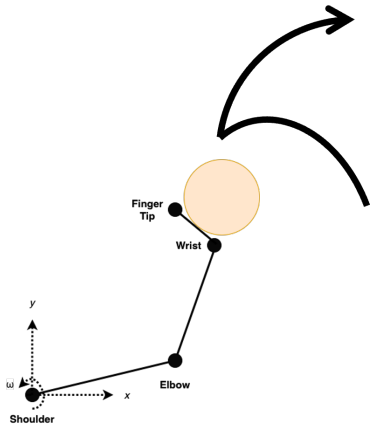


Figure 2.5: Two-motion shot

Motor Coordination

Bernstein's research in the field of motor control uncovers how coordinated motion is achieved through mastery of the degrees of freedom of movement [18]. When individuals learn a new motor skill, they tend to pause at convenient degrees of freedom to simplify the motion. As individuals gain more experience with performing the motion, they "release"

and exploit the degrees of freedom to gain proficiency [18]. Pertaining to basketball, a novice player will appear rigid when learning to shoot a free throw. With practice, the elbow and wrist joints within the shooting arm move more fluidly and efficiently, transferring energy in a proximal-to-distal direction rather than through simultaneous movement.

Kinematics and Kinetics of the Shooting Arm

Okubo and Hubbard developed a kinematic [153] and kinetic [154] model of the shooting arm. Using a 2D model with three rigid planar links and rotational joints to represent the upper arm, forearm, and hand, the authors studied the dependence of optimal release conditions and the corresponding arm motion and joint torques using a dynamic simulation model. At the moment of release, they found that the shoulder contributes to the vertical component of the ball release velocity, while the elbow and wrist contribute to the horizontal velocity and backspin. The authors stressed that the optimal release condition consists of a combination of shoulder, elbow, and wrist angular velocities (see Figure 2.6).

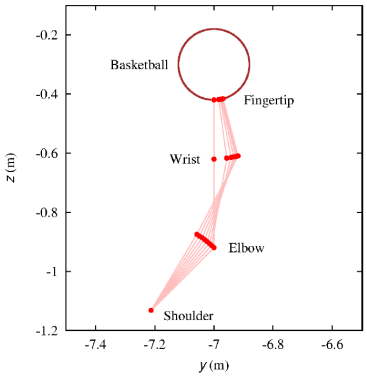


Figure 2.6: Set of arm joint positions at release [153]

As the angular velocity of the upper arm increases, the angular velocity of the forearm increases in the opposite direction. Logically, a large wrist angular velocity is required if the shoulder and elbow angular velocities are small. At ball release, the torques of the joints are positive, the shoulder angular acceleration is positive, while the elbow and wrist accelerations are negative.

2.1.3 Free Throws in Wheelchair Basketball

Laurie Malone was among the first researchers to investigate the mechanics of the wheelchair basketball free throw across different classifications. During the 1994 Gold Cup World Wheelchair Basketball Championships, Malone et al. [128] studied the performance of free throws from Class 1.0 athletes by describing the kinematic differences between successful and unsuccessful shots. Athletes in this study employed either a pushing motion or a throwing motion and leaned against the backrest during their free throws. Regardless of the shooting form, the authors noticed that elbow and shoulder velocities were greater during successful attempts.

Using the same data from the 1994 Gold Cup, Malone et al. [127] expanded their analysis of the free throw by considering how release height, angle, and speed differed among athletes of different classifications. Labeling Class 1 and Class 2 as the "lower class" and Class 3 and Class 4 as the "upper class," significant differences were observed in the release parameters during the free throw; lower-class athletes shoot from a lower release height and use a higher release angle and speed compared to higher-class athletes. Malone et al. [131] also investigated the different types of outcomes of a free throw beyond a *make* (swish) or a *miss*. This dichotomous system does not adequately provide feedback to the shooter because free throw outcomes exist across a spectrum (clean swish (CS), backboard (BB), back rim (BR), front rim (FR), subsequent rim bounce (R), miss (M), success (H)) [131]. Shooters can better assess their shooting form by understanding the type of miss or success in their attempt.

Malone et al. [129], Goosey-Tolfrey and Butterworth [71], and Prvulović [166] examined the biomechanics of the shooting form across different classifications. Using video capture data of free throws, they examined shoulder angle, elbow angle, trunk inclination angle, shoulder axis (torso rotation), angular displacements, and velocities of the major joints. Since athletes with a lower classification tend to use a steeper release angle and a higher release speed, they must produce greater force in their shooting arm. This is accomplished through a more flexed elbow position during the preparatory phase of the free throw to increase elbow range of motion and generate the necessary impulse during arm elevation [166]. Additionally, higher angular velocities at the shoulder and elbow were observed. Higher-class athletes have a more efficient shot due to the engagement of their trunk muscles and a higher release height [71].

2.2 2D Human Pose Estimation

Human pose estimation is a challenging computer vision task that determines the geometric posture and, the movement of human body parts from visual data. Pose estimation infers the configuration of an articulated human body, which consists of joints and rigid body parts. Classical approaches to human pose estimation used pictorial structures [11][96] and human body models to represent the spatial correlations of body parts through graphical/tree models [199].

Early works also incorporated hand-crafted edge and contour detectors, along with histogram of oriented gradients (HOG) features [212]. In general, human pose models based on these techniques suffered from a lack of high accuracy and generalization [146].

The rise and success of deep learning in computer vision have outperformed and replaced classical methods in various tasks that pose estimation is based on, including image classification [109], semantic segmentation [122], and object detection [168]. There are various applications of pose estimation, including action recognition [125], person detection [116], and human tracking [84]. Pose estimation is also implemented in virtual reality, human-computer interaction, surveillance [7], self-driving [66], sports motion analysis, animations, and video games [109].

Despite the significant progress in pose estimation, challenges remain inherent due to the nature of the problem. Occlusion of joints due to complex poses, crowding, or intricate body configurations can skew the model's prediction of a joint position or misidentify body parts. Similarly, spatial interference caused by clothing, lighting, and varying camera angles can also negatively impact pose estimation. Figure 2.7 highlights the challenges of pose estimation.



Figure 2.7: Challenges of human pose estimation in images or videos. Examples are from the Max Planck Institute for Informatics (MPII) dataset [10]

Organization of Sections

The following section will outline the recent advances in 2D human pose estimation (HPE) using deep learning and its applications in sports. Section 2.2.1 will highlight the common deep learning architectures in HPE. Finally, section 2.2.2 will summarize the popular datasets and evaluation metrics followed by a comparison of results of the different methods.

2.2.1 Deep Learning Approaches

Before deep learning, traditional human modeling approaches used hand-engineered features to extract information of body parts and joints and utilized graphical tree structures to approximate the configuration of limbs between connecting joints for pose estimation [47]. However, these methods could not generalize well to complex poses due to the lack of spatial inference between body parts that were further away. Fortunately, recent advances of deep learning, primarily the success of convolutional neural networks, better examines the configuration of a person on a larger scale and derive localized relationships between adjacent joints and the body parts resulting in significantly improved performance.

There are two categories of deep learning-based single person pose estimation: *regression* methods and body part *detection* methods (Figure 2.8). Regression methods map input images directly to joint coordinates or human body model parameters [191], but are less robust due to the nonlinear nature of the problem. Detection methods predict body

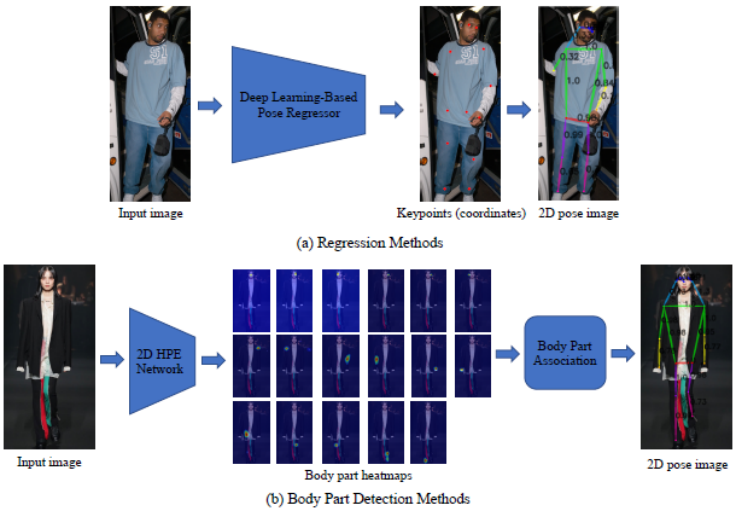


Figure 2.8: (a) Regression method maps body joint coordinates from the input image. (b) Detection method predicts joint coordinates from representative heatmaps [220].

part or joint locations [33, 149], typically using image patches or heatmaps [90, 149], and are more robust as they learn from dense pixel information.

Regression Methods

Regression methods directly map input images to joint coordinates, offering computational efficiency for pose estimation. One of the earliest breakthroughs, DeepPose [191], introduced convolutional neural networks (CNNs), such as AlexNet, to predict joint coordinates. Its cascaded architecture used intermediate supervision to iteratively refine predictions, setting the stage for modern regression-based approaches.

To enhance regression accuracy, several methods incorporated additional contextual or structural information. For instance, the Iterative Error Feedback (IEF) [27] leveraged feedback loops with GoogLeNet [183] to iteratively refine joint localization. Sun et al. [182] proposed a bone-based representation that integrated body structure into pose estimation, moving beyond traditional joint-based representations.

Simple Baselines, proposed by Xiao et al. [205], significantly simplified regression-based architectures while achieving competitive accuracy. Using a ResNet backbone followed by a few deconvolutional layers, it generated high-resolution heatmaps efficiently, as shown in Figure 2.9. Unlike complex architectures like the stacked hourglass, Simple Baselines

demonstrated that simplicity and elegance can rival sophisticated multi-stage networks in accuracy and computational efficiency.

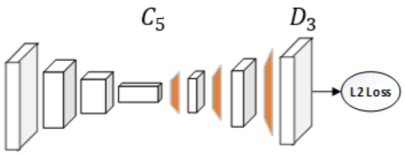


Figure 2.9: Architecture of Simple Baselines: ResNet backbone with deconvolution layers for heatmap generation [205].

Regression methods also adopted multi-task learning to enhance feature representation. Luvizon et al. [125] combined pose estimation with action recognition, while Li et al. [117] introduced a framework that jointly predicted joint coordinates and body parts from image patches. These innovations addressed challenges like generalizability and feature sparsity in regression-based pose estimation.

Despite their strengths, regression methods often struggle with the nonlinear complexity of pose estimation. However, their computational efficiency makes them appealing for real-time applications.

Detection Methods

Detection-based methods predict joint coordinates through heatmaps, which offer superior spatial localization and supervision. Heatmaps encode the probability of joint positions in pixel space, enabling precise keypoint estimation. Ground-truth heatmaps are generated using 2D Gaussian distributions centered at annotated joint locations, and loss functions like Mean Squared Error (MSE) are minimized during training.

One of the most influential detection-based architectures, the Stacked Hourglass Network [149], introduced a bottom-up and top-down structure to process multi-scale features iteratively. Each stage refines joint predictions through pooling and upsampling, effectively capturing spatial relationships. The High-Resolution Net (HRNet) [181] further advanced detection-based methods by maintaining high-resolution representations throughout the network. Unlike architectures that downsample features to low resolution and attempt to recover them, HRNet used parallel high-to-low-resolution subnetworks with multi-scale fusions, achieving state-of-the-art performance without requiring intermediate heatmap supervision.

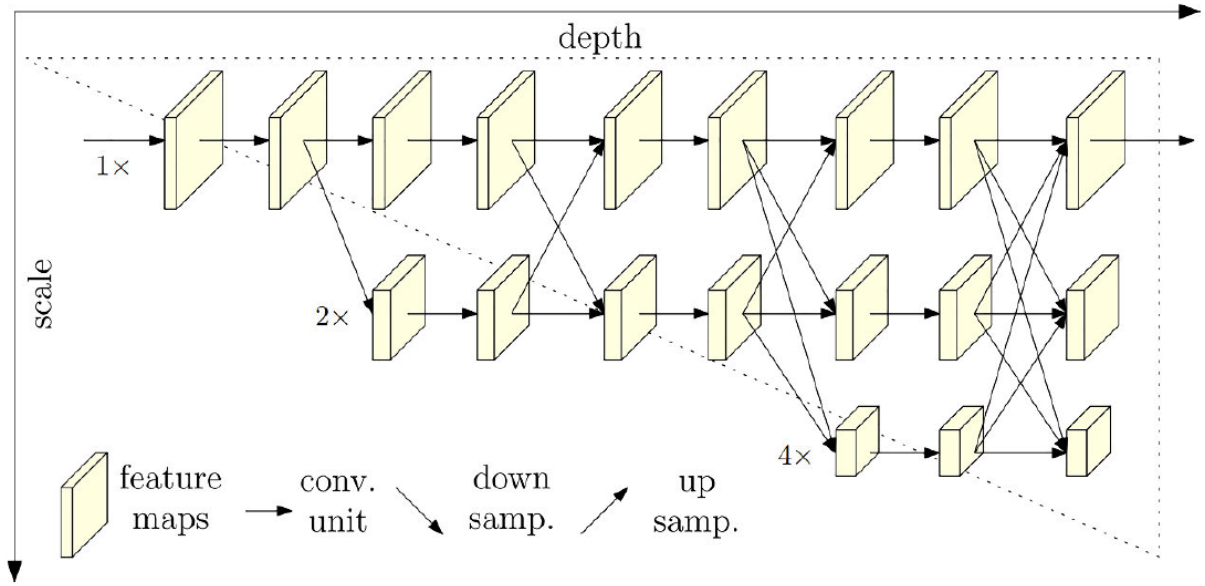


Figure 2.10: Architecture of HRNet [181]

To address quantization errors in heatmap resolution, Zheng et al. [216] introduced DarkPose, a distribution-aware coordinate representation technique. DarkPose enhances spatial accuracy by refining resolution recovery, making it a vital preprocessing step for many heatmap-based models.

Multi-stage frameworks, such as Convolutional Pose Machines (CPMs) [201], iteratively refined heatmap predictions to improve spatial consistency. CPMs used sequential modules with large receptive fields, mitigating vanishing gradient issues while capturing long-range dependencies. Similarly, Newell et al. [149] employed skip connections in their hourglass architecture to preserve spatial information across multiple resolutions.

Neural architecture search (NAS) has also been applied to detection-based methods. **ViPNAS** [207] automated the design of pose estimation architectures by optimizing both accuracy and efficiency. This approach produced networks that outperformed manually designed counterparts, demonstrating the promise of algorithmic architecture optimization.

While detection-based methods excel in spatial localization, they face challenges related to computational cost and memory usage for high-resolution heatmaps. However, innovations like HRNet, DarkPose, and ViPNAS have significantly advanced the field, balancing accuracy and efficiency.

Summary of 2D Pose Estimation

Deep learning has revolutionized 2D pose estimation, with regression and detection methods offering distinct advantages. Regression models excel in efficiency, making them suitable for real-time applications, but their performance can be limited by nonlinear problem complexity. Detection frameworks, particularly heatmap-based approaches, provide superior spatial accuracy but are computationally intensive.

Challenges remain, including improving real-time implementation, handling unusual poses, and addressing dataset imbalances. Architectures like HRNet and Simple Baselines highlight the importance of balancing complexity, accuracy, and computational efficiency, while techniques like DarkPose mitigate bottlenecks in resolution recovery. Future advancements may include robust data augmentation, lightweight architectures, and further applications of NAS to optimize pose estimation frameworks.

2.2.2 Datasets and Evaluations

Shown below is an unabbreviated overview of the datasets.

Image-based Datasets

- Frames Labeled in Cinema (FLIC)
- Leeds Sports Pose (LSP)
- Max Planck Institute for Informatics (MPII)
- Microsoft Common Objects in Context (COCO)[1]
- AI Challenger Human Keypoint Detection (AIC-HKD)
- CrowdPose

Video-based Datasets

- Penn Action

[1]The traditional COCO dataset has 17 keypoints annotations for each person in an image. COCO *Whole-Body* consists of 133 keypoints for each person. The pose estimation models in this thesis were trained on the latter.

- Joint-annotated Human Motion Database (J-HMDB)
- PoseTrack

Figure 2.11 summarizes the current benchmark datasets and evaluation metrics in 2D pose estimation.

Image-based datasets							
Name	Year	Single-Person /Multi-Person	Joints	Size			Evaluation protocol
				Train	Val	Test	
FLIC	2013	Single	10	5k	-	1k	PCP&PCK
FLIC-full	2013	Single	10	20k	-	-	PCP&PCK
FLIC-plus	2014	Single	10	17k	-	-	PCP&PCK
MPII	2014	Single	16	29k	-	12k	PCPm/PCKh
		Multiple	16	3.8k	-	1.7k	mAp
LSP	2010	Single	14	1k	-	1k	PCP
LSP-extended	2011	Single	14	10k	-	-	PCP
COCO2016	2016	Multiple	17	45k	22k	80k	AP
COCO2017	2017	Multiple	17	64k	2.7k	40k	AP
AIC-HKD	2017	Multiple	14	210k	30k	60k	AP
CrowdPose	2019	Multiple	14	10k	2k	8k	mAP
Video-based datasets							
Name	Year	Single-Person /Multi-Person	Joints	Size			Evaluation protocol
				Train	Val	Test	
Penn Action	2013	Single	13	1k	-	1k	-
J-HMDB	2013	Single	15	0.6k	-	0.3k	-
PoseTrack	2017	Multiple	15	292	50	208	mAP

Figure 2.11: 2D single person pose estimation datasets and metrics [220]

Common Evaluation Metrics

Percentage of Correct Parts (PCP):
This metric measures the detection rate of limbs. Correct detection occurs if the distance between the two predicted joint locations and the true limb joints is less than half of the limb length [191].

Percentage of Detected Joints (PDJ):
This metric deems the correct prediction of a joint if the distance between the true joint location and the predicted joint location is within a certain fraction of the torso diameter.

Percentage of Correct Keypoints (PCK):
This metric measures the accuracy of the location of the body joint. A correct prediction

occurs if the distance between the true joint location and the predicted joint location is within a certain fraction of a side length of an external rectangle of ground truth body joints [151].

PCKh:
Similar to PCK, but the matching threshold is based on 50% of the head length. Joint detection is correct if the predicted location is within this threshold from the true joint location.

Procrustes Aligned Mean Per-Joint Position Error (PA-MPJPE):
This metric evaluates the accuracy of human pose estimation algorithms by measuring the average Euclidean distance between predicted and ground-truth joint positions after aligning them using Procrustes analysis, which involves translation, rotation, and scaling transformations to minimize differences [9].

Chapter 3

Biomechanical Model of Basketball Shooting

The basketball free throw is predominantly shot with one arm. A free throw shooter holds the ball in their preferred hand, bringing it beside their head or near their forehead, and quickly extends their elbow while flexing their wrist to release the ball from the tips of their index and/or middle finger. The other arm stabilizes the ball just before release. Previous research suggests that a shooter's shoulder, elbow, wrist, and ball must follow a vertical plane in line with the basketball net for an efficient shooting motion [102]. Shooters must release the ball at their highest release point with ample backspin, while minimizing lateral deviation of their hand and ball.

3.1 Rationale for 2D Analysis

The biomechanical analysis presented in this thesis employs a 2D sagittal plane approach. While 3D analysis provides comprehensive spatial information and has been recommended by some researchers [75], several compelling theoretical and practical reasons support the use of 2D analysis for wheelchair basketball free throw biomechanics.

Biomechanical Justification

The basketball free throw is characterized by biomechanical constraints that naturally favor planar motion. Skilled shooters maintain alignment of the shoulder, elbow, wrist,

and ball in a vertical plane directed toward the basket [101]. This alignment principle, often referred to as the "plane of motion" in coaching literature, suggests that deviations from the sagittal plane are typically associated with shooting errors rather than optimal technique [54].

For wheelchair basketball specifically, the seated position and strapping system further constrain the athlete's movement. Lower-class athletes (1.0-2.5) have limited trunk mobility due to their disability, while even higher-class athletes are restricted by the chair's backrest and their secured position [130]. These physical constraints reinforce the predominance of sagittal plane motion during free throw execution.

Mathematical Modeling Considerations

From a biomechanical modeling perspective, the 2D approach offers several advantages:

1. **Reduced Complexity**: The equations of motion and inverse dynamics calculations are significantly simplified in 2D, allowing for clearer interpretation of joint torques and angular velocities.
2. **Validated Precedent**: Previous biomechanical studies of basketball shooting have successfully used 2D models to derive meaningful insights about optimal release conditions and joint coordination patterns [153].
3. **Parameter Reduction**: A 2D model reduces the number of degrees of freedom from 7 (for a 3D arm model) to 3, making the system more tractable for analysis while retaining the essential biomechanical information.

Alignment with Coaching Practice

Traditional coaching methodologies for free throw analysis have long relied on sagittal plane observation. Coaches position themselves to the side of the shooter to assess key technical elements including release height, elbow alignment, and follow-through mechanics [107]. The primary coaching cues—"elbow under the ball," "high arc," "full extension," and "consistent release point"—all refer to movements and positions observable from the side view.

Scope and Limitations

This methodological choice acknowledges that 3D analysis could provide additional insights, particularly regarding ball spin axes and subtle compensatory movements. However, the 2D approach captures the primary biomechanical variables relevant to free throw performance: joint angles, angular velocities, and torques in the dominant plane of motion.

The trade-off between comprehensive 3D data and 2D analysis was carefully considered. Given that the ultimate goal is to provide actionable biomechanical feedback for wheelchair basketball athletes and coaches, the 2D approach offers an appropriate level of detail while maintaining analytical tractability. This aligns with recent trends in sports biomechanics that emphasize practical applicability over laboratory-confined comprehensiveness [41].

The following sections present the biomechanical model designed within these 2D constraints, demonstrating how meaningful insights about shooting mechanics can be derived from sagittal plane analysis.

3.2 Design of Model

The shooting arm model is similar to the 2D three-segment model proposed by Schwark et. al [173] that was used to optimize the shooting torques and the release velocity of the basketball using Powell's optimization search. The XYZ frame with unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ is fixed to the ground with the XY plane lying horizontally and the Z axis pointing up. The arm model is comprised of 3 segments and 3 rotational joints to represent the upper arm, forearm, and hand and the shoulder, elbow, and wrist joints. The arm and ball move in a vertical plane, as illustrated in Figure 3.1. Points S, E, W, T denote the shoulder, elbow, wrist joint centres and the fingertip, respectively. The segments U, F, and H represent the upper arm, forearm and hand, respectively. The lengths of each segment are L_U, L_F and L_H. The centre of mass of the segments are indicated by the points U^*, F^* and H^*. The centre of the ball is shown as B^*. Joint angles can be used to determine the position of the fingertip: θ_S is the angle between the horizontal plane and the upper arm, θ_E is the angle between the upper arm and the forearm, and θ_W is the angle between the forearm and the hand. Arm segment positions can also be determined using the counterclockwise angular displacement (θ_S, θ_F, θ_H) of the segment from the horizontal plane.

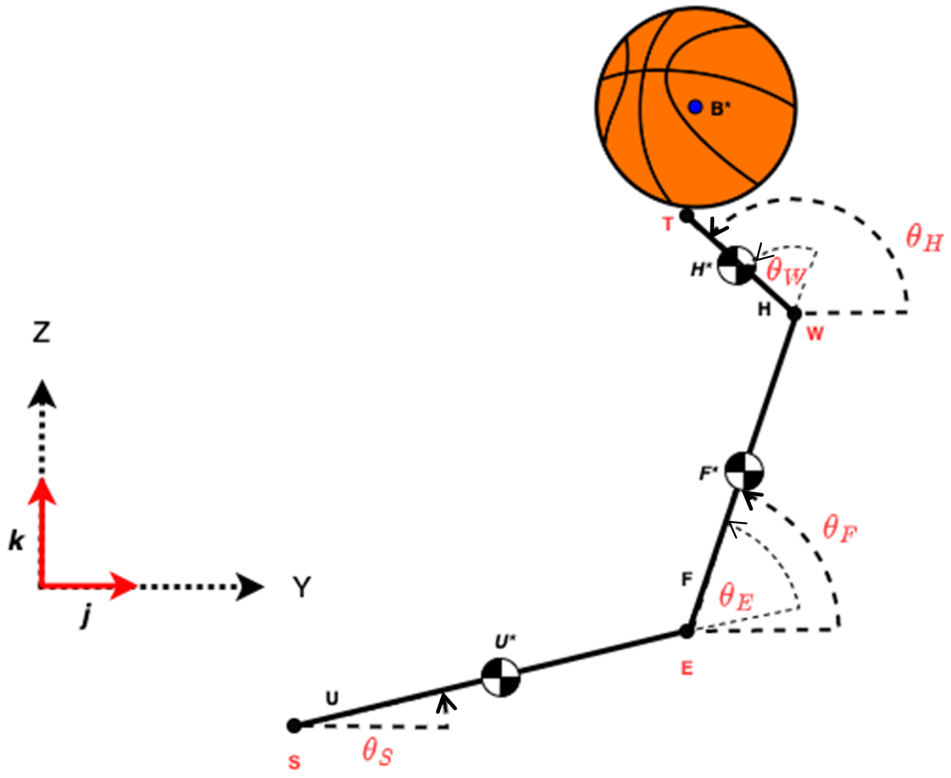


Figure 3.1: Geometric representation of the shooting arm during a free throw

Assumptions

The biomechanical analysis of the shooting arm during a free throw is based on some assumptions:

- Motion occurs in a 2D sagittal plane. This assumption is justified by biomechanical studies showing that proficient free throw shooters maintain their shooting arm movement primarily in a single vertical plane [101], and deviations from this plane are generally associated with shooting errors. For wheelchair basketball athletes, the fixed seating position and potential trunk limitations further constrain movement to the sagittal plane.
- Shoulder joint centre (S) is fixed (horizontal and vertical velocity is 0).
- Hand segment (H) is a rigid body (joints within the hand are fixed).

- Centre of mass of the ball (B^*) is coincident with the fingertip (T).
- There is no slipping between the contact point of the hand and the ball (T) during release.

If the basketball is assumed to be secured to the hand until release, the velocity of the contact point between the hand and the ball (T) can approximate the velocity of the basketball. Note that the angular velocity of the basketball is not modeled.

Implications of Excluding Trunk Dynamics

While the three-segment model captures the primary mechanics of the shooting arm, the exclusion of trunk dynamics represents a significant simplification with important biomechanical implications. Research has demonstrated that trunk function correlates positively with wheelchair basketball player classification [156], with trunk control being the primary determinant in the International Wheelchair Basketball Federation (IWBF) classification system [86].

The trunk serves as the foundational segment for force generation in wheelchair basketball shooting. Lower-class players (1.0-2.5) with limited trunk control must compensate through increased upper limb torques, as observed in this study's findings. Malone et al. [130] reported that wheelchair basketball players classified as 1 or 2 on the IWBF system find it difficult to generate force from the shoulder since their unstable trunk cannot support the reaction forces. This compensation mechanism results in:

1. **Altered Force Generation Patterns**: Without trunk contribution, the shoulder must act as both the primary force generator and stabilizer, leading to the significantly higher shoulder torques observed in class 1.0 athletes.
2. **Reduced Mechanical Efficiency**: The inability to utilize trunk momentum requires athletes to rely solely on upper extremity strength, resulting in less efficient shooting mechanics and potentially lower release heights [72].
3. **Limited Coaching Insights**: As noted by the principle "what gets measured, matters," the absence of trunk metrics may inadvertently direct coaching attention away from trunk stabilization strategies that could benefit even lower-class athletes within their functional limitations.

Despite these limitations, the focused analysis of upper limb mechanics provides valuable insights into the compensatory strategies employed by athletes with different trunk function levels. The significantly higher joint torques observed in lower-class athletes quantify the biomechanical cost of limited trunk involvement, information that remains clinically relevant even without explicit trunk modeling.

3.3 Kinematic Analysis

Rigid body kinematic equations are used to derive the velocity of the hand of a three-link planar arm during a free throw. The velocity of the hand can indicate the release speed and angle of the basketball during release based on the aforementioned assumptions. Inputs to the model are the angles and angular velocities of the shoulder, wrist and elbow joints. The horizontal ($v_{T'_y}$) and vertical ($v_{T'_z}$) component of the velocity of the fingertip is expressed as a function of the angles ($\theta_S, \theta_E, \theta_W$) and angular velocities ($\dot{\theta}_S, \dot{\theta}_E, \dot{\theta}_W$) of the shoulder, elbow and wrist, and the lengths of the upper arm, forearm and centre of mass of the hand from the wrist ($L_U, L_F, L_{H_{CM}}$):

$$v_{T'_y} = -\dot{\theta}_S L_U sin(\theta_S) - \dot{\theta}_F L_F sin(\theta_F) - (\dot{\theta}_H) L_{H_{CM}} sin(\theta_H) \tag{3.1}$$

$$v_{T'_z} = \dot{\theta}_S L_U cos(\theta_S) + \dot{\theta}_F L_F cos(\theta_F) + (\dot{\theta}_H) L_{H_{CM}} cos(\theta_H) \tag{3.2}$$

where:

$$\theta_F = \theta_S + \theta_E \tag{3.3}$$

$$\theta_H = \theta_S + \theta_E + \theta_W \tag{3.4}$$

3.4 Inverse Dynamics

The precise coordination of muscle contractions within the shooting arm results in a successful free throw. Fatigue has adverse effects on the consistency of the muscle control and is detrimental to the shooting form [197]. Net joint torques, which serve as a surrogate measure of muscular effort, can be calculated using an inverse dynamics analysis. This technique applies the Newton-Euler equations of motion to the shooting arm model, using the arm segment kinematics and anthropometric parameters as inputs. By using inverse dynamics analysis, coaches and athletes can gain insights into the muscular contributions and coordination required for shooting, allowing for better training and performance optimization [154].

The joint reaction forces ($\boldsymbol{F}_S$, $\boldsymbol{F}_E$ and $\boldsymbol{F}_W$), position of centre of mass of each segment ($\boldsymbol{r}$) and joint torques ($\boldsymbol{\tau}_S$, $\boldsymbol{\tau}_E$ and $\boldsymbol{\tau}_W$) are illustrated in Figure 3.2. Note that the centre of mass of the basketball is assumed to be coincident with the fingertip (T).

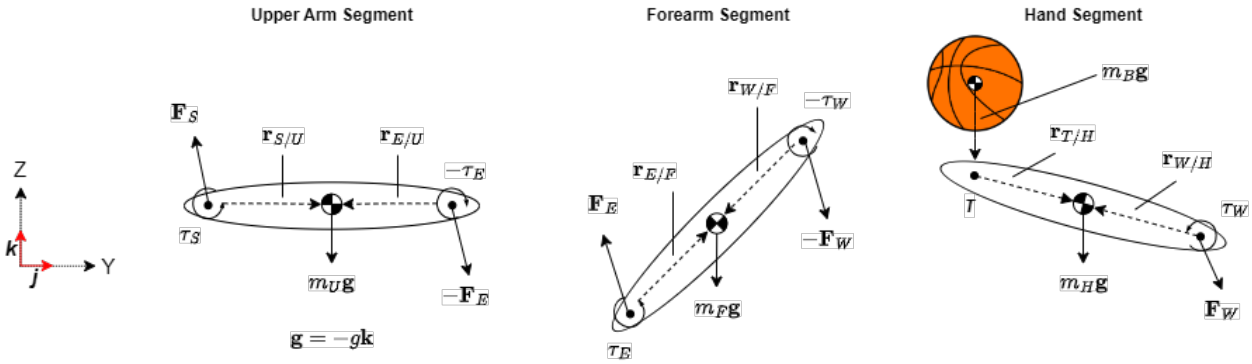


Figure 3.2: Free body diagram of arm segments

The anthropometrics, segment length, segment mass, and moment of inertia (I_U, I_F, I_H) of the shooting arm were calculated using data from de Leva [48]. The inputs to the inverse dynamic models are the accelerations of the centre of masses of the segments ($\boldsymbol{a}_U$, $\boldsymbol{a}_F$, $\boldsymbol{a}_H$), derived using 2D rigid body kinematics, and angular accelerations of the segments ($\boldsymbol{\alpha}_U$, $\boldsymbol{\alpha}_F$, $\boldsymbol{\alpha}_H$), the second derivative of angular positions. The recursive Newton-Euler approach was used to derive the joint torques.

Accelerations of Joint Centres / Segment Centre of Mass:

$$\mathbf{a}_C = \mathbf{a}_A + \boldsymbol{\alpha} \times \boldsymbol{r}_{C/A} - \dot{\theta}^2 \boldsymbol{r}_{C/A} \tag{3.5}$$

($\boldsymbol{a}_C$ represents the acceleration at the centre of mass of the limb segment while $\boldsymbol{a}_A$ represents the acceleration at the proximal joint centre of the same limb segment)

Torque Calculations via Recursive Newton-Euler Inverse Dynamics:

$$\mathbf{F}_W + (m_H + m_B)\mathbf{g} = (m_H)\mathbf{a}_{H_{CM}} + (m_B)\mathbf{a}_B \quad (I_H + I_B)\boldsymbol{\alpha}_H = -\boldsymbol{r}_{W/H} \times \boldsymbol{F}_W - \boldsymbol{r}_{T/H} \times (-m_B\mathbf{g}) + \boldsymbol{\tau}_W \tag{3.6}$$

$$\mathbf{F}_E - \mathbf{F}_W + m_F\mathbf{g} = m_F\mathbf{a}_{F_{CM}} \qquad I_F\boldsymbol{\alpha}_F = -\boldsymbol{r}_{E/F} \times \boldsymbol{F}_E - \boldsymbol{r}_{W/F} \times (-\boldsymbol{F}_W) + \boldsymbol{\tau}_E - \boldsymbol{\tau}_W \tag{3.7}$$

$$\mathbf{F}_S - \mathbf{F}_E + m_U\mathbf{g} = m_U\mathbf{a}_{U_{CM}} \qquad I_U\boldsymbol{\alpha}_U = -\boldsymbol{r}_{S/U} \times \boldsymbol{F}_S - \boldsymbol{r}_{E/U} \times (-\boldsymbol{F}_E) + \boldsymbol{\tau}_S - \boldsymbol{\tau}_E \tag{3.8}$$

Segment	**Length (m)**	**Centre of Mass (m)**	**Mass (kg)**	**Moment of Inertia xx (kgm^2)**
Upper Arm	0.282	0.162	2.03	0.0125
Forearm	0.269	0.123	1.22	0.0060
Hand	0.189	0.086	0.46	0.0046
Basketball	-	-	0.625	0.0059

Table 3.1: Anthropometric parameters for torque calculations based on a male (M; 75.0 kg, 1.741 m) [48]

3.5 Validation

Validation of the torque calculations using the Newton-Euler inverse dynamics approach was performed using the MapleSim 2022 software. MapleSim is a multi-domain modeling and simulation environment that uses a combination of symbolic and numeric computation techniques to analyse multibody systems. In addition to generating system equations, MapleSim provides simulation code that can be implemented in hardware.

Instrumentation and Software

The motion of the shooting arm of a wheelchair basketball athlete in the sagittal plane was recorded using an iPhone XS smartphone, with a 1080p resolution, 60 Hz rear-facing

video camera. The (y,z) coordinates of the joint centres of the arm in each frame of the video were manually annotated using the Physlets Tracker software [53]. The joint centre coordinates were used to calculate the kinematics of the arm segments. The coordinates for calculation of a particular joint form a triangle. The law of cosines is applied to determine the joint angles after calculating the Euclidean distances between all coordinates. Arm segment kinematics serve as inputs to the Newton-Euler and MapleSim inverse dynamics models for joint torque calculations.

Data Processing

Inverse dynamics calculations require the differentiation of the displacement of the joint angles to obtain angular velocities and accelerations. Derivatives were calculated using the central difference differentiation [175]. Differentiation of the raw signal induces amplification of high-frequency noise, distorting downstream analysis. Attenuation of this noise was accomplished using a 2^{nd}-order low-pass Butterworth filter with a cutoff frequency of 6 Hz.

Results

As shown in Figure 3.3, the Newton-Euler derived joint torques are compared to the MapleSim model and are nearly identical, implying that the equations for the derived model are correct.

3.6 Discussion

The plots in Figure 3.3 depict the joint torques of the shooting arm from the transition phase to the release point of a successful free throw from a wheelchair basketball participant. The vertical line separates the transition phase from the launch phase at the start of elbow extension. As expected, the magnitude of the shoulder torques are greater than those of the elbow and wrist because the shoulder is supporting the combined load of the entire arm and the basketball. Based on the coordinate system shown in Figure 3.1, the shoulder torque is positive because the arm rotates in a counterclockwise direction. The elbow and wrist torque are negative because the forearm and hand rotate in a clockwise direction to facilitate the extension of the arm [154].

The shoulder torque increases during the transition phase and launch phase. This occurs due to the increase in the length of the moment arm about the shoulder joint centre

as the shooting arm extends to release the ball, combined with the activation of muscle to raise the upper arm. This is consistent with the findings from Okubo and Hubbard [154], who demonstrate that shoulder torques are at their peak during the release of the ball. The magnitude of the elbow torque decreases during the end of the transition phase and at the beginning of the launch phase because the moment arm about the elbow joint decreases; the forearm is perpendicular to the ground and the positions of the centre of masses of the basketball and the hand lie directly above the elbow joint, resulting in a smaller moment arm about the elbow. The magnitude of elbow torque increases during elbow extension due to muscle activation and the increasing moment arm, aligning with Schwark's joint torque optimization model for a free throw in wheelchair basketball [173]. Finally, the magnitude of the wrist torque increases from the transition phase to the launch phase due to the rapid flexion of the wrist to propel the basketball away from the shooting arm and provide backspin [102].

Incorporating inverse dynamics into training is valuable for coaches and athletes seeking to optimize performance, prevent injuries, and enhance overall athletic development. The joint torques of the shooting arm play a crucial role in influencing the shooting mechanics and accuracy of a free throw. While various release conditions can lead to successful free throws [16], the shooting form that utilizes the least joint torque is preferred because it minimizes energy expenditure and chance of injury [197]. Moreover, inverse dynamics can reveal weaknesses in specific muscle groups. Inverse dynamics aids in designing individualized training programs that target specific muscle groups and enhance arm motion coordination and provides a quantifiable approach to measure the performance of the shooting form.

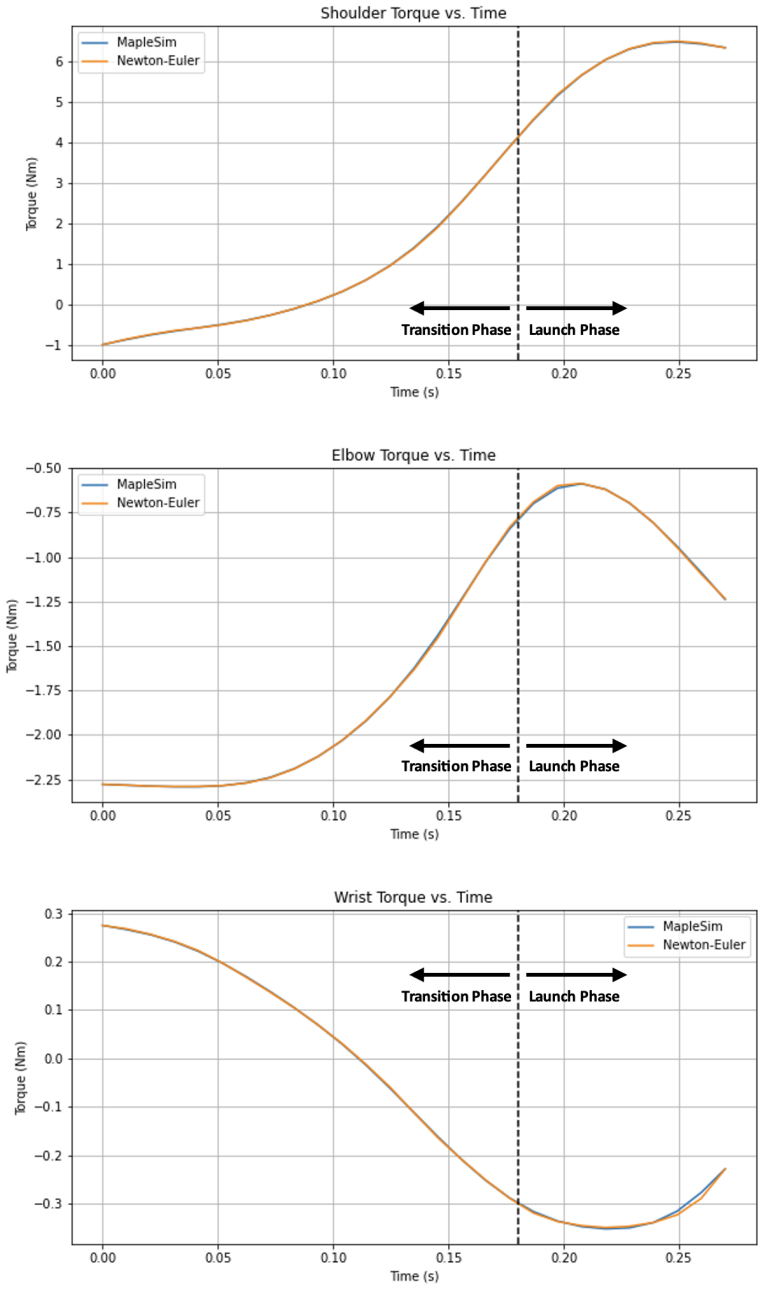


Figure 3.3: Joint torque validation of Newton-Euler approach against MapleSim for the shoulder, elbow and wrist. Vertical line separates the transition phase from the launch phase.

Chapter 4

Pose Estimation for Basketball Shooting

This thesis proposes a pose estimator that uses the convenient and accessible form-factor of a smartphone to provide biomechanical information of the shooting arm during a basketball shot. The mobile application, *ShotPose*, leverages the smartphone's camera and hardware resources to capture videos of free throw shooting and employ deep learning algorithms for processing. By combining pose estimation with kinematic and dynamic analysis, as shown in Figure 4.1, this free throw shooting mobile application aims to inform athletes and coaches about the efficiency of shooting form and develop training strategies in any environment.

4.1 Computer Vision using Edge Computing

4.1.1 Deployment of Computer Vision using Core ML

The Core ML framework from Apple allows developers to integrate trained machine learning (ML) models into mobile apps for devices running iOS 11 or later. It provides support for a wide range of ML models, including neural networks, tree ensembles, support vector machines, and generalized linear models. Typically, ML models are pre-trained on a computer and then saved on the smartphone for accessibility by the app. Core ML also offers on-device training capabilities for devices with iOS 13 or later.

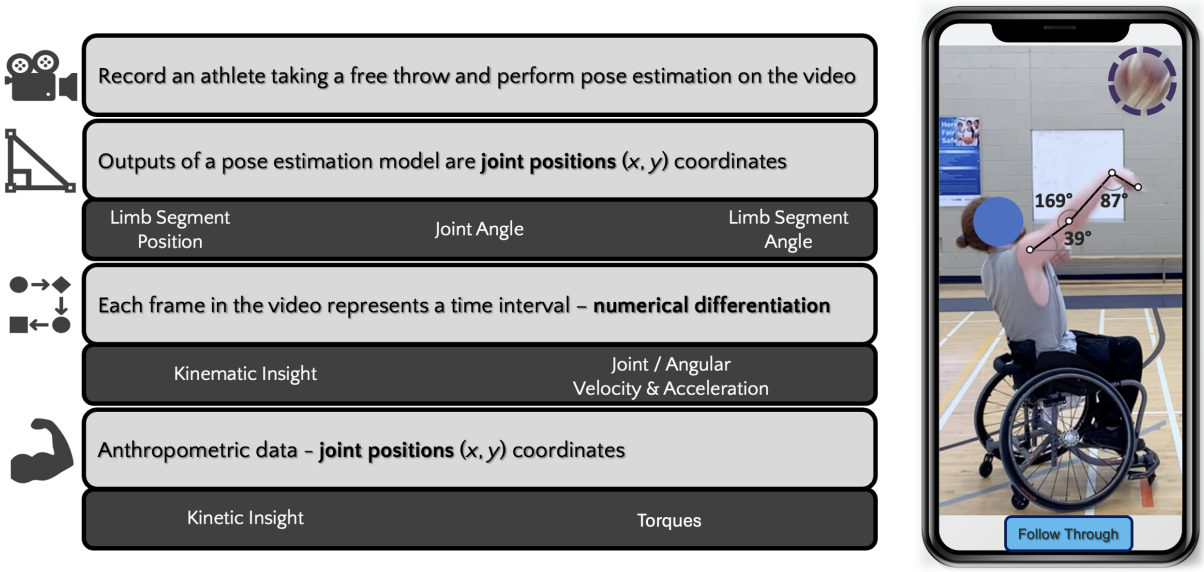


Figure 4.1: Overview of the pose estimation algorithm

Through Core ML's Vision API, developers can harness the power of smartphone cameras to perform various computer vision tasks, such as object detection, tracking, face/landmark detection, and pose estimation. For pose estimation, Core ML provides a pre-trained model based on *PoseNet*. However, it should be noted that PoseNet [100] is not suitable for *ShotPose* as it cannot detect the joints in the hand, which are necessary for calculating the wrist angle. State-of-the-art pose estimation models like PoseNet are trained and benchmarked on image datasets such as COCO or CrowdPose, where each person instance in an image is annotated with 17 keypoints (excluding hand joints).

Core ML supports the deployment of custom neural networks developed using PyTorch or TensorFlow. The *MMPOSE* [43] open-source repository provides several pre-trained pose estimation models that can detect 133 keypoints, including 21 keypoints for each hand. These models were trained on the COCO-Wholebody dataset [95], which contain more than $200K$ images and $250K$ person instances with labeled keypoints.

4.1.2 MMPose Pose Estimation Models

MMPose offers several pre-trained models for different applications of pose estimation. *HRNet* [181], *Simple Baselines* [205], *ViPNAS* [207] and *HRNet Dark+* [216] were trained on COCO-Wholebody (capable of detecting joints within the hand) and were investigated

during the development of *ShotPose*. All four models were initially trained on the COCO dataset, and later fine-tuned on the COCO-Wholebody dataset for improved joint detection accuracy. Each model was trained on the COCO *train2017* dataset with COCO-Wholebody annotations (133 keypoints), including $57K$ images and $150K$ person instances. Evaluation was performed on the *val2017* dataset and *test-dev2017* dataset containing 5000 images and $20K$, respectively. Detailed experimental configuration of each model (data augmentation techniques, type of deep learning optimizer, learning rate, number of epochs and batch size) can be found on MMPose Wholebody 2D Keypoint Model Zoo webpage [138].

The design features of each model are briefly discussed below.

HRNet

HRNet employs a multi-resolution fusion strategy that preserves fine-grained details while effectively integrating information from different levels of the network. This strategy involves constructing parallel sub-networks that process input images at different resolutions and then progressively fuse their features to obtain a final high-resolution representation. This approach allows the model to capture both low-resolution global context and high-resolution local details simultaneously [181].

Simple Baselines

The Simple Baselines architecture for pose estimation leverages ResNet as the backbone network for image feature extraction [78]. In this approach, several deconvolutional layers are added after the last convolutional stage of the ResNet architecture to generate heatmaps. These deconvolutional layers have batch normalization and ReLU activation, and each has 256 filters with a 4x4 kernel and a stride of 2. A 1x1 convolutional layer is added at the end to produce predicted heatmaps for all keypoints. The model is trained using Mean Squared Error (MSE) loss between the predicted heatmaps and the ground truth heatmaps, which are created using 2D Gaussians centered on the joint's ground truth locations .

This architecture is chosen for its simplicity and efficiency in generating heatmaps from deep and low-resolution features, similar to the approach used in Mask R-CNN [77]. The Simple Baselines method combines upsampling and convolutional parameters into deconvolutional layers in a straightforward manner, avoiding the use of skip layer connections. This simplicity makes it easy to implement and effective for various pose estimation tasks. [205].

ViPNAS

The ViPNAS algorithm addresses the need for efficient video pose estimation by employing a neural architecture search method that considers both spatial and temporal levels, resulting in a better trade-off between accuracy and real-time efficiency in comparison to the traditional heatmap-based approach. Neural Architecture Search (NAS) is a method used to automatically search and design optimal neural network architectures using strategies such as reinforcement learning, evolutionary learning or gradient-based methods.

The ViPNAS framework has two levels of search space: spatial-level and temporal-level. In the spatial level, ViPNAS designs a search space with five dimensions: network depth, width, kernel size, group number, and attentions. By exploring various combinations within these dimensions, the algorithm aims to find an architecture that balances accuracy and efficiency. The temporal-level search space includes a series of temporal feature fusions to optimize the total accuracy and speed across multiple video frames [207].

HRNet Dark+ (DarkPose)

The Distribution-Aware coordinate Representation of Keypoint (DARK) method tackles the computation inefficiency in pose estimation by focusing on the resolution recovery process during heatmap generation. To address this issue, the authors propose an improved encoding process that generates more precise heatmap distributions, effectively reducing quantization errors commonly found in standard pose estimation models. Additionally, the DARK algorithm enhances the heatmap decoding step by introducing a method based on the distribution structure of the predicted heatmap. This method aims to infer the underlying maximum activation accurately. Serving as a preprocessing step, DARK is applied before utilizing keypoint-generating pose estimation models. In the development of the proposed mobile application discussed in this thesis, DARK is used in conjunction with HRNet, and is referred to as DarkPose [216].

Model Metrics

Table 4.1 compares the performance of the four pose estimation models on the Coco-Wholebody dataset. The average precision (AP) of the hand and body keypoints are shown. AP is a widely used evaluation metric in pose estimation tasks, specifically to assess the accuracy of a model in localizing body keypoints on a human subject. A pose estimation model predicts the positions of various keypoints like the shoulder, elbow or

Model	Body AP	Hand AP	Latency (ms)	Size (MB)
HRNet	0.722	0.587	65.2	243
Simple Baselines	0.703	0.559	47.3	263
ViPNAS	0.643	0.410	27.8	30
DarkPose	0.742	0.602	76.4	243

Table 4.1: Performance metrics and sizes of MMPose models on the COCO-Wholebody dataset

wrist joint centres. To compute AP, several steps are followed. First, the model predicts the keypoint locations for each instance in the test dataset. Next, the Intersection over Union (IoU) is calculated for each predicted keypoint, which measures the alignment between the predicted and ground-truth keypoints. The IoU values are then used to create a precision-recall curve, illustrating the trade-off between precision (accuracy of predictions) and recall (correct predictions made). By calculating the average precision (AP) as the area under this curve, the average accuracy of the model across all recall values is determined. Higher AP values indicate better performance and more accurate localization of keypoints by the pose estimation model. The average precision at the IoU threshold α (AP^{α}) is the area under the precision-recall curve *p(r)* for confidence thresholds from 0 to 1:

$$AP^{\alpha} = \int_0^1 p(r)dr \tag{4.1}$$

The latency of each model was measured on an NVIDIA GeForce RTX 2070 GPU running Ubuntu 20.04 by recording the time required to process an input image. This measurement indicates the processing time a user would experience when using the pose estimation model for their recorded video. Among the models tested, DarkPose proved to be the most accurate, though it is memory-intensive and slower compared to ViPNAS. In contrast, NAS models are generally lightweight and fast, but they compromise on accuracy. Given that precise detection of joint centers and fingertips is crucial for the biomechanical analysis of the shooting arm, the DarkPose model was deemed most suitable for *ShotPose*.

4.2 Implementation on an iPhone

Shown in Figure 4.2 is a wireframe of the app.

4.2.1 Walkthrough of *ShotPose*

ShotPose enables athletes and coaches to apply pose estimation to recorded videos of free throw shots stored on the device. To capture the shooting arm in the sagittal plane, the mobile device should be positioned beside the shooter, aligned with the free throw line. For optimal results, it is recommended to mount the device on a tripod.

Once a video is recorded and saved in the device's memory, *ShotPose* can access it and trim it to the desired time frames, such as the loading phase or the launch phase of the shot. This allows users to analyze specific parts of the video. The pose estimation model will then process the trimmed video and generate a new video with visualizations and biomechanical information. The resulting video will be saved back to the gallery for further review. An example of this process is depicted in Figures 4.3 and 4.4.

Biomechanical Analysis
The following metrics can be shown on the visualization video:

- Joint angles and angular speeds
- Joint torques

Refer to kinematic and inverse dynamic model equations in Chapter 3.

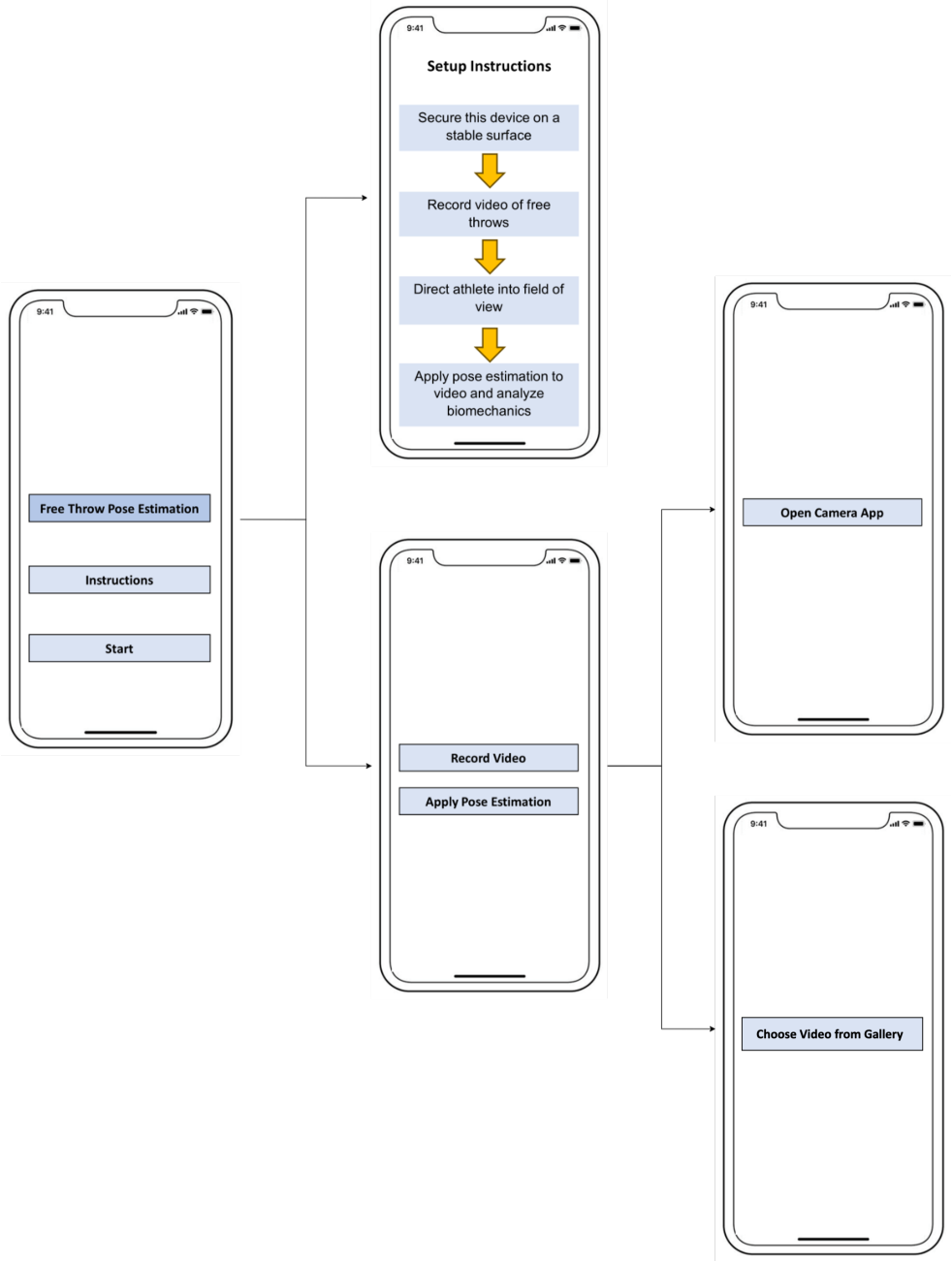


Figure 4.2: Wireframe of *ShotPose* in the IPhone XS template



Figure 4.3: Frame of a video before pose estimation



Figure 4.4: Pose estimation result with joint torques ($N \cdot m$) and limb segment angles (deg)

Refer to Figures 3.1 and 3.2 for definitions of positive joint angles and torques.

4.2.2 Data Processing

Refer to Section 3.5 for the method of data processing for differentiation and noise attenuation for inverse dynamics calculations.

4.3 Experiments with *ShotPose*

The purpose of this experiment is to determine the accuracy of the pose estimation capability of *ShotPose* on an iOS device by comparing it to the marker-based Vicon motion capture system. Additionally, this study aims to investigate the efficiency of the one-motion shot versus the two-motion shot, and compare the biomechanical differences between an able-bodied free throw shooter sitting in a wheelchair and a low-class wheelchair basketball free throw shooter.

4.3.1 Method

Subjects

Fifteen healthy adults (five females; 11 males; 24 ± 3.2 years; height 1.74 ± 0.08 m; mass 73.26 ± 16.73 kg) were recruited from the University of Waterloo. One male subject (28 years; height 1.80 m; mass 72 kg) is a T4 complete paraplegic and is confined to a wheelchair. He is a member of the Canadian Men's Wheelchair Basketball team and has a classification of 1.0 according to the International Wheelchair Basketball Federation (IWBF).

The able-bodied participants were divided into two groups based on their basketball experience: recreational and competitive. The recreational group consisted of eleven participants with diverse shooting forms, while the competitive group included four male participants with consistent shooting forms. All participants were right-handed and had no other physical conditions or limitations that would prevent them from shooting basketballs.

Protocol and Instrumentation

Before the study began, the participants shot 20 free throws to warmup. The study required participants to shoot free throws until they made 10 swishes (ball enters the net without making contact with the rim).

Able-bodied participants were instructed to perform the shooting exercise in two positions: while standing and while sitting on a basketball wheelchair, in order to minimize leg power generation during the shot and replicate the shooting style of a wheelchair basketball athlete with a 4.5 IWBF classification. However, three female and two male participants lacked the necessary arm strength to execute a free throw, so only their standing attempts were recorded.

The motion of the shooting arm was captured using the Vicon motion capture system (Vicon, Oxford, UK), which consisted of eight infrared cameras sampling at 100 Hz. The data was recorded using Vicon Nexus 2.14 mobile. Additionally, an iPhone XS smartphone with a 1080p resolution and a 60 Hz rear-facing video camera was used. The smartphone was accessed by the free throw app, developed using XCode 14.1 and running on iOS 16.

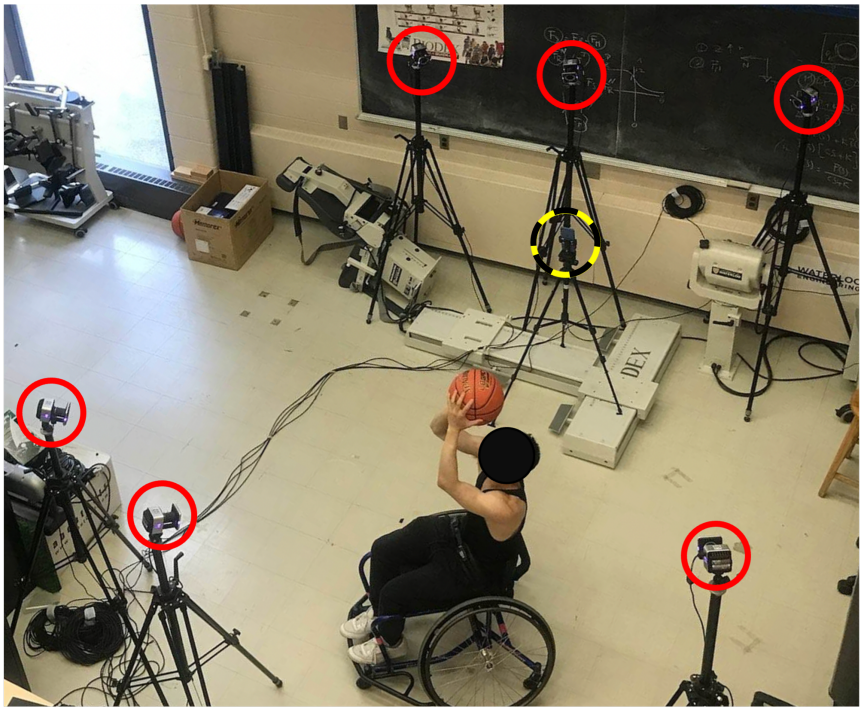


Figure 4.5: Red (solid): Vicon cameras, Yellow/black (dashed): iPhone XS

System Calibration

A two-step calibration process of the Vicon cameras was performed prior to data collection. The first step involved the motion of an LED calibration wand within the entire 3D space the cameras would record. This is to ensure that the eight Vicon cameras have a synchronized global fixed axis. Next, the wand was placed on the ground so that the cameras recognized the position of the floor as the origin.

The second step calibrates the anatomical landmarks of the subject, represented by the Vicon optical markers within the global fixed axis, and ensures visibility of the markers to a majority of the eight cameras. Palpation was performed to identify joint centers of the shoulder, elbow, and the wrist during the placement of the markers. The optical marker on the hand could not be placed on the index finger due to insufficient surface area for adhesion, and was instead placed beside the metacarpal of the pinky finger. A short recording of the subject was obtained to classify the position of the markers as ball joints on the Nexus software. The Vicon cameras were calibrated so that the wheelchair (or any reflective surface) were masked during the recording.

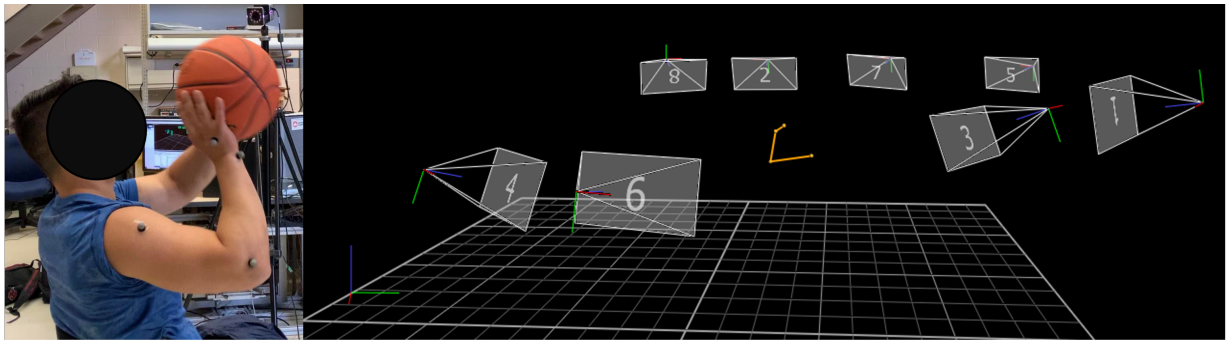


Figure 4.6: Left: Placement of the optical markers on the shooting arm. Right: Visualization of the shooting arm and position of the Vicon cameras in the Nexus software.

Limitations of the VICON Setup

It should be noted that the VICON marker placement protocol used in this study differs from the best practices recommended by the International Society of Biomechanics (ISB) [204]. The ISB standards recommend using clusters of at least three non-collinear markers per segment to define anatomical coordinate systems and minimize the effects of soft tissue artifacts.

The current study's approach of using single markers placed on palpated joint centers has several limitations:

1. **Soft Tissue Artifacts**: Single markers are susceptible to skin movement relative to the underlying bone, particularly during dynamic movements. This can introduce errors of 10-30mm in marker position [114].

2. **Joint Center Estimation**: Palpation-based marker placement is subject to inter-operator variability and may not accurately represent the true instantaneous center of rotation for each joint.

3. **2D Projection**: Converting 3D marker positions to 2D coordinates for comparison with pose estimation eliminates out-of-plane information and may introduce projection errors.

4. **Lack of Segment Tracking**: Without multiple markers per segment, the true 6DOF motion of each limb segment cannot be determined, limiting the analysis to simplified joint angles.

These limitations mean that the VICON data, while more precise than pose estimation in terms of marker tracking, may not represent the true gold standard for skeletal kinematics. The comparison should therefore be interpreted as a relative assessment between two measurement approaches, each with inherent limitations.

Data Processing

Temporal synchronization between the 100 Hz Vicon system and 60 Hz iPhone pose estimation required careful consideration due to the different sampling rates and potential system latencies. Following established biomechanical synchronization protocols [105], an event-based synchronization approach was employed using a distinct kinematic event visible in both systems. The initiation of arm movement from the static preparatory position served as the synchronization anchor point, as this transition from zero to non-zero angular velocity provides a clear, unambiguous temporal landmark [65].

The synchronization process involved: (1) identifying the frame where shoulder angular velocity first exceeded a threshold of 5°/s in both systems, indicating movement onset; (2) aligning these frames as the common temporal reference point (t=0); and (3) accounting for the temporal offset between systems, which ranged from 0-16.7 ms due to the asynchronous nature of the camera triggers. This event-based approach is preferred over simple frame matching as it accounts for variable system delays that can range from 4-235 ms between trials [105].

Gaps in optical marker detection within the Vicon data were addressed using spline interpolation. The same spline interpolation technique was also applied to upsample the pose estimation data collected from the iPhone at 60 Hz to match the 100 Hz frequency of the Vicon motion capture data. This upsampling process was necessary to ensure a proper comparison between the iPhone pose estimation data and the Vicon data. The raw data obtained from the Vicon system consisted of joint locations represented by 3D coordinates. However, since *ShotPose* performs 2D pose estimation, only the coordinates corresponding to the motion of the arm in the global sagittal plane (as derived from Vicon) were used to calculate joint angles. As described in Section 3.5, a low-pass Butterworth filter is applied before differentiation to further refine the data.

4.3.2 Comparison to Marker-based System

Methodological Considerations for Validation

The validation comparison between *ShotPose* and VICON must be interpreted within the context of each system's limitations. While VICON is often considered the gold standard for motion capture, the simplified marker protocol used in this study introduces several sources of error:

1. **Marker placement errors**: Studies have shown that inconsistent marker placement can contribute up to 5 degrees of error in joint angle calculations [74, 174]. The use of single markers amplifies this issue.
2. **Soft tissue artifacts**: During the dynamic movements of free throw shooting, skin-mounted markers can move relative to the underlying bone by 10-30mm, particularly at the shoulder and elbow [113].
3. **2D projection artifacts**: The process of extracting 2D coordinates from 3D marker data for comparison with pose estimation may introduce additional errors, particularly if the movement deviates from the assumed sagittal plane.

Given these considerations, the reported RMSE values should be viewed as a comparison between two imperfect measurement systems rather than an absolute measure of pose estimation error. The true skeletal kinematics likely fall somewhere between the two measurements, and the acceptable error threshold for practical coaching applications may be higher than traditional biomechanical research standards.

Joint Angle Accuracy

Biomechanical analysis of the shooting arm stems from the shoulder, elbow and wrist joint angles. Therefore, the accuracy of the pose estimation models used in *ShotPose* can be evaluated by comparing its calculated joint angles with those obtained from the Vicon motion capture system. *ShotPose* and the Vicon system estimate the coordinates of the shoulder, elbow, wrist and finger joint centres. The coordinates for calculation of a particular joint form a triangle. The law of cosines is applied to determine the joint angles after calculating the Euclidean distances between all coordinates (derivation shown below).

The coordinates of the joint centers for the shoulder, elbow, wrist, and fingertip in the 2D plane are defined as follows:

- Shoulder (s): (y_S, z_S)
- Elbow (E): (y_E, z_E)
- Wrist (W): (y_W, z_W)
- Fingertip (T): (y_T, z_T)

The shoulder angle (θ_S) is the angle between the upper arm and the horizontal axis, as shown in 4.7. Using the coordinates of the shoulder and elbow joint centres:

$$\theta_S = \arctan\left(\frac{z_E - z_S}{y_E - y_S}\right) \tag{4.2}$$

The elbow angle (θ_E) is the angle between the upper arm and the forearm. Using the coordinates of the shoulder, elbow, and wrist joint centres, we can use the Law of Cosines:

Let a, b, and c be the lengths of the segments as follows:

$$a = \sqrt{(y_W - y_S)^2 + (z_W - z_S)^2} \quad \text{(length from S to W)} \tag{4.3}$$
$$b = \sqrt{(y_E - y_S)^2 + (z_E - z_S)^2} \quad \text{(length from S to E)} \tag{4.4}$$
$$c = \sqrt{(y_W - y_E)^2 + (z_W - z_E)^2} \quad \text{(length from E to W)} \tag{4.5}$$

The elbow angle is then given by:

$$\theta_E = 180^\circ - \arccos\left(\frac{b^2 + c^2 - a^2}{2bc}\right) \tag{4.6}$$

The wrist angle is the angle between the forearm and the hand. Using the coordinates of the elbow, wrist, and fingertip joint centres, we can use the Law of Cosines again:

Let d, e, and c be the lengths of the segments as follows:

$$d = \sqrt{(y_T - y_E)^2 + (z_T - z_E)^2} \quad \text{(length from E to T)} \tag{4.7}$$
$$e = \sqrt{(y_T - y_W)^2 + (z_T - z_W)^2} \quad \text{(length from W to T)} \tag{4.8}$$
$$c = \sqrt{(y_W - y_E)^2 + (z_W - z_E)^2} \quad \text{(length from E to W, as calculated previously)} \tag{4.9}$$

The wrist angle is then given by:

$$\theta_W = \arccos\left(\frac{c^2 + e^2 - d^2}{2ce}\right) \tag{4.10}$$

Figure 4.7 shows how the joint angles are defined.

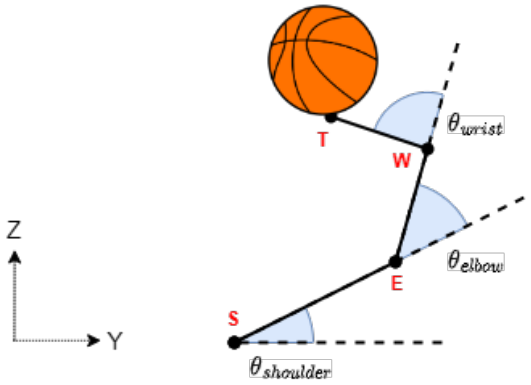


Figure 4.7: Definition of joint angles

Root mean squared errors (RMSEs) between the joint angles of each of the pose estimation models deployed in *ShotPose* ($\hat{y}$) and the Vicon system (y) for N joint angles were calculated using the following formula:

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{\sum_{i=0}^{N-1}(y_i - \hat{y}_i)^2}{N}} \tag{4.11}$$

The results are shown in Table 4.2, and indicate that DarkPose is the most accurate of the models. This is expected since DarkPose has the highest AP.[1]

	HRNet	**Simple Baselines**	**ViPNAS**	**DarkPose**
Shoulder	8.18°	7.52°	8.29°	6.85°
Elbow	17.22°	17.63°	19.36°	16.43°
Wrist	68.71°	72.84°	85.58°	63.57°

Table 4.2: Joint angle RMSE of *ShotPose* vs. Vicon

[1] All pose estimation analysis/plots presented using DarkPose

As expected, the joint angle RMSEs increase with distal joints because the wrist and finger move quickly in the frame of the video, and appear blurry, as shown in Figure 4.8. Motion blur hinders the pose estimation model's ability to accurately detect the precise location of keypoints within the pixels of the image. The imprecise prediction of the wrist joint due to motion blurring contribute to the difference in elbow angle calculations between the Vicon system and the pose estimation model as shown in Figure 4.10. A phone camera with a frame rate of at least 120 fps would mitigate the effects of motion blur.

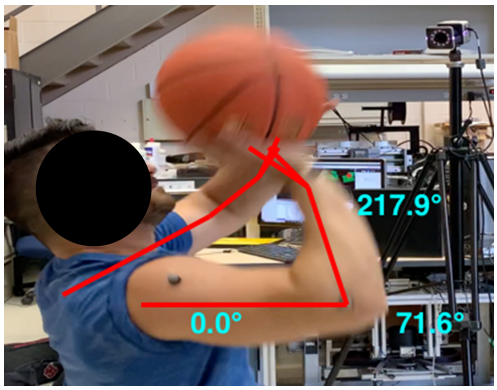


Figure 4.8: Motion blur appearing in a frame of a free throw video

RMSEs of the wrist angle, shown in Table 4.2, indicate that the accuracy of the pose estimation model is poor when compared to the Vicon system. However, as shown in Figure 4.10, the high RMSE value are a result of a consistent offset. This is because the Vicon system's measurement of the wrist angle is not based on the true position of the finger, rather, on the position of the marker placed near the metacarpal of the pinky, shown in Figure 4.9. The wrist angle calculated by *ShotPose* is based on the position of the finger tip.

Biomechanical Validity of the Wrist Angle Measurement

The wrist angle calculation in this study represents a composite measurement of wrist and hand orientation rather than isolated wrist flexion/extension. This approach, while simplified compared to multi-segment hand models, remains biomechanically valid and interpretable for basketball shooting analysis based on several considerations:

1. **Functional Unity During Shooting**: Throughout the basketball free throw, the hand maintains a relatively stable configuration with the ball resting on the palm and

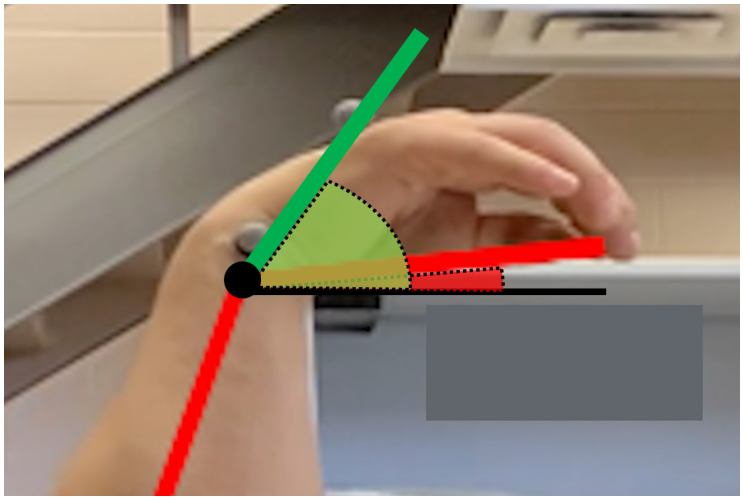


Figure 4.9: Wrist angle discrepancy between optical marker placement (green solid line) and detection of finger (red solid line) by the pose estimation model from *ShotPose*

fingertips. Research on basketball shooting mechanics demonstrates that the hand-ball system functions as a quasi-rigid unit until the final moments of release [23]. The primary movement occurs at the radiocarpal joint (true wrist), with minimal metacarpophalangeal or interphalangeal flexion until ball release.

2. **Biomechanical Precedent**: Previous biomechanical analyses of basketball shooting have successfully used similar simplified hand models. Elliott [54] and Okubo & Hubbard [154] treated the hand-ball complex as a single segment when analyzing shooting mechanics, recognizing that the functional outcome (ball trajectory) depends on the combined wrist-hand orientation rather than isolated joint movements.

3. **Measurement Practicality**: The composite wrist-hand angle captures the functionally relevant parameter for shooting analysis—the orientation of the ball release vector. Since the basketball obscures individual finger joints during most of the shooting motion, and finger flexion contributes minimally until the final release instant, the simplified measurement provides the most reliable and repeatable metric for comparative analysis.

4. **Clinical Relevance**: In rehabilitation and coaching contexts, practitioners typically assess and train the wrist-hand complex as a functional unit during shooting. The composite angle measurement aligns with how coaches observe and correct shooting form, making it more practically applicable than isolated joint measurements.

While acknowledging that metacarpophalangeal flexion contributes to backspin generation during release, this occurs over a brief time window (typically ¡50ms) that is difficult to capture reliably at 60Hz sampling rates. The composite wrist angle measurement therefore provides a robust, interpretable metric that captures the primary biomechanical function while avoiding the measurement uncertainties associated with tracking individual finger segments during rapid movements.

This methodological choice is further validated by the consistent offset observed between Vicon and pose estimation measurements (Figure 4.9), suggesting that both systems are capturing related but slightly different aspects of hand orientation—a systematic difference rather than random error.

Joint Coordinate Accuracy

The utilisation of markerless motion capture (pose estimation) in biomechanics and rehab applications have been bottlenecked due to the lack of precision in the detection of joints [42], [160]. Moreover, Hellsten *et al.* [79] designed a simulation to visualize how the imprecision in joint detection is significantly compounded in the accuracy of joint angle calculations. Their results suggest that the probability of measuring the true angle of the joint within a $\pm 5^{\circ}$ margin is only 25% when the prediction of the joint coordinates are within a 10 cm radius. The error margin in the joint coordinate estimation must be at most 2.5 cm to guarantee a maximum error margin of $\pm 5^{\circ}$ in joint angle calculation.

The mean joint coordinate error of the shoulder, elbow and wrist joint centres between the Vicon optical markers and pose estimation predictions are 14 cm, 12 cm and 29 cm, respectively. The pose estimation model accuracy of the distal joints decreases drastically during the launch phase of the shot due to motion blurring. Based on Hellsten *et. al*'s simulation, the joint coordinate error of the pose estimation model exist beyond the 10 cm radius. This suggests that the probability of measuring the true angle of the joints within a $\pm 5^{\circ}$ margin would be less than 25% for the shoulder angle, and substantially less for the elbow and wrist angles.

4.3.3 Assessing Wheelchair Basketball Free Throws using Motion Capture

Motion capture provides a quantified approach to studying the unique shooting form of each athlete, which is especially valuable in wheelchair basketball as it allows coaches to characterize the mechanics of the shooting arm based on shot type and athlete classification.

While *ShotPose* represents a promising markerless motion capture modality for accessible biomechanical analysis, the current implementation exhibits joint angle errors that exceed clinically acceptable thresholds for certain applications.

Biomechanical literature suggests that joint angle root mean square errors (RMSE) below 5° are considered excellent for clinical and sports applications, while errors between 5-10° are deemed acceptable for many movement analysis scenarios [169, 195]. Recent markerless motion capture systems have achieved grand mean errors of 3.85-6.3° RMSE across various movements [177, 147]. In the current implementation, *ShotPose* demonstrates acceptable accuracy for shoulder joint angles (RMSE = 6.85°), marginal accuracy for elbow angles (RMSE = 16.43°), and substantial errors for wrist angles (RMSE = 63.57°). The elevated errors, particularly at distal joints, are primarily attributed to motion blur during rapid shooting movements and the challenges of tracking smaller anatomical landmarks at 60 Hz sampling rates.

These error magnitudes, especially for the elbow and wrist joints which are critical in basketball shooting mechanics, preclude definitive biomechanical conclusions from *ShotPose* data alone. Following established practices in biomechanical validation studies [98], when markerless system errors exceed acceptable thresholds, marker-based motion capture data should be used for detailed biomechanical analysis. Therefore, this section employs Vicon motion capture data to investigate the biomechanics of wheelchair basketball free throws, establishing reference patterns that can inform future iterations of *ShotPose* when enhanced pose estimation models or higher frame rate cameras reduce measurement errors to clinically acceptable levels.

The following analysis examines biomechanical differences between one-motion versus two-motion shooting techniques and compares shooting mechanics across athlete classifications (Class 1.0 vs. Class 4.5) using VICON motion capture data.

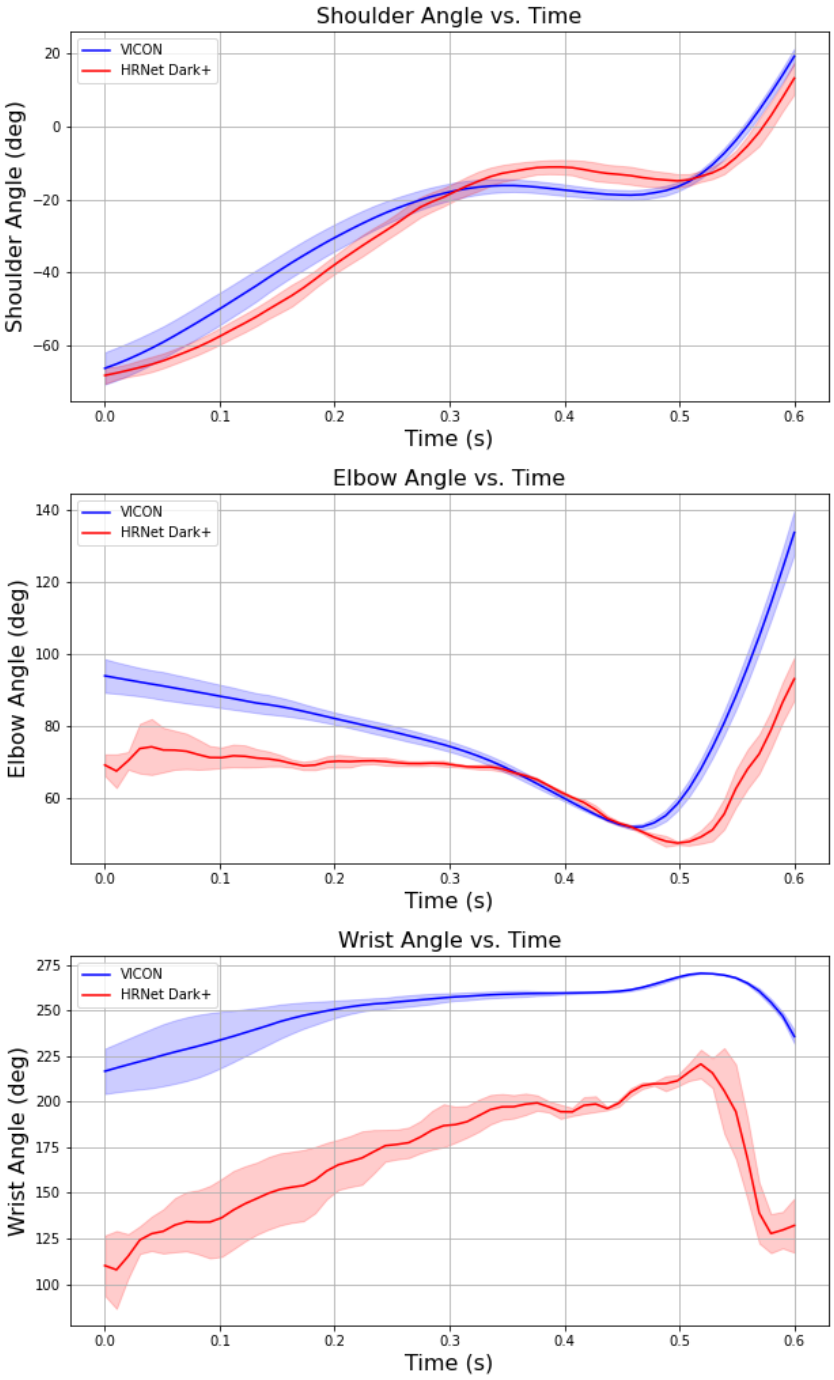


Figure 4.10: Mean joint angles: Vicon vs. Free throw app. (Shaded region represents ±1 std dev)

One-Motion Shot vs. Two-Motion Shot

During the experiment, participants exhibited two primary shooting motions. Among the recreational group participants, those who stood employed a one-motion shot technique, loading the basketball from the hip. On the other hand, among participants who were competitive in able-bodied basketball, two individuals utilized a one-motion form, while the remaining two employed a two-motion form. In the case of able-bodied participants shooting from a basketball wheelchair, all of them employed a one-motion form.

Figure 4.11 displays a comparison of the shooting arm's biomechanics between the standing one-motion and two-motion shots, as analyzed through pose estimation. The plots in the figure depict the transition phase and launch phase of successful shots. The distinction between a one-motion shot versus a two-motion shot is determined empirically based on the movement of the shooting arm. In the case of a one-motion shot, the upper arm's angular velocity gradually increases throughout the shot's progression. Conversely, in a two-motion shot, a distinct "pause" is observed as the upper arm's angular velocity reaches 0 deg/s, indicating the initiation of the launch phase. Furthermore, the extension of the elbow is less pronounced in the one-motion shot compared to the two-motion shot. Lastly, it is worth noting that the angular velocity of the hand during wrist flexion is notably higher in the two-motion shot.

A coach can determine the efficiency of the shooting form by analyzing the dynamics of the arm. Joint torques and powers can be calculated using motion capture data and anthropometric information. Figure 4.11 indicates that shoulder and elbow torques are higher in a two-motion shot. Furthermore, a one-motion shot generates more power at the shoulder and elbow joints. The differences in wrist torque and power between the one-motion and two-motion shots appear negligible.

The results suggest that the one-motion shot is more efficient compared to the two-motion variant, as it generates lower joint torque. The high torque and low power in the two-motion shot can be attributed to the loss of momentum when the arm and ball come to a stop at the shoulder before the launch phase. To achieve the desired launch conditions, the shooter must exert additional force to accelerate the ball from rest at the shoulder. This requires a rapid extension of the elbow, resulting in high elbow torque. In contrast, in a one-motion shot, the ball moves over a longer distance without interruption. With uninterrupted motion between the transition and launch phases, it becomes easier to generate the necessary ball speed at release for a successful free throw. This could explain why able-bodied participants exclusively used a one-motion shot when shooting from a basketball wheelchair.

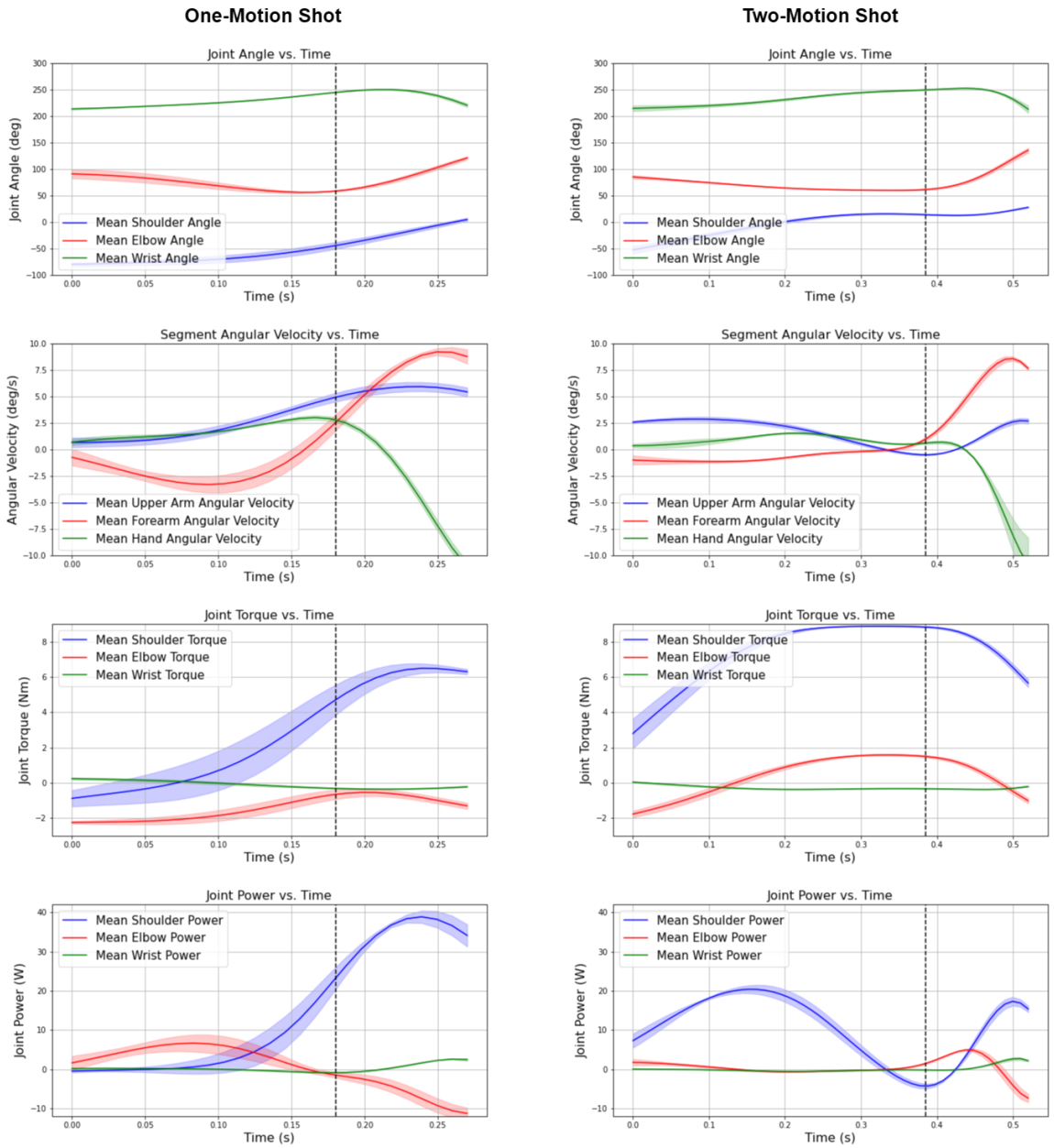


Figure 4.11: One-motion shot (left column) vs. Two-motion shot (right column). Vertical line (start of elbow extension) separates the transition phase from the launch phase. (Shaded region represents ±1 std dev)

Class 1.0 vs. Class 4.5

A similar experiment was conducted to investigate the difference in shooting form between a class 1.0 wheelchair basketball player and a class 4.5 wheelchair basketball player. In this study, athletes with 4.5 classification were represented by competitive able-bodied basketball players sitting on a wheelchair. Able-bodied participants belonging to the "competitive" category were chosen for this study because they maintain a consistent technique, similar to the national level wheelchair basketball participant. In addition to the difference in functional ability, the class 1 participant releases the basketball from a lower height (1.65 m) and the class 4.5 participants release the basketball from an average height of (1.78 m).

Figure 4.12 reveals distinct differences in shooting form between the class 1 participant and the class 4.5 participants. The shooting form of the class 1 participant resembles a two-motion shot, with the upper arm's angular velocity reaching 0 deg/s at the beginning of the launch phase. In contrast, the class 4.5 participants adopt a one-motion shooting form.

At release, the class 1 participant exhibits higher shoulder and elbow angles, along with faster angular velocities of the upper arm, forearm, and hand segments. Additionally, the class 1 participant generates higher shoulder and elbow joint torques compared to the class 4.5 participants. These results are expected for two reasons:

1. Since the class 1 participant is shooting from a lower release height, the arm must be positioned in a way to provide a higher release angle (more arc) to increase the margin for a successful shot.

2. The class 4.5 athletes extend their trunk during the transition phase of the shot to provide additional force to the basketball. The class 1 participant cannot engage his trunk muscles due to the severity of his paraplegia. Thus, he must rely solely on the rapid extension of his arm to provide the required force, resulting in production of higher joint torques.

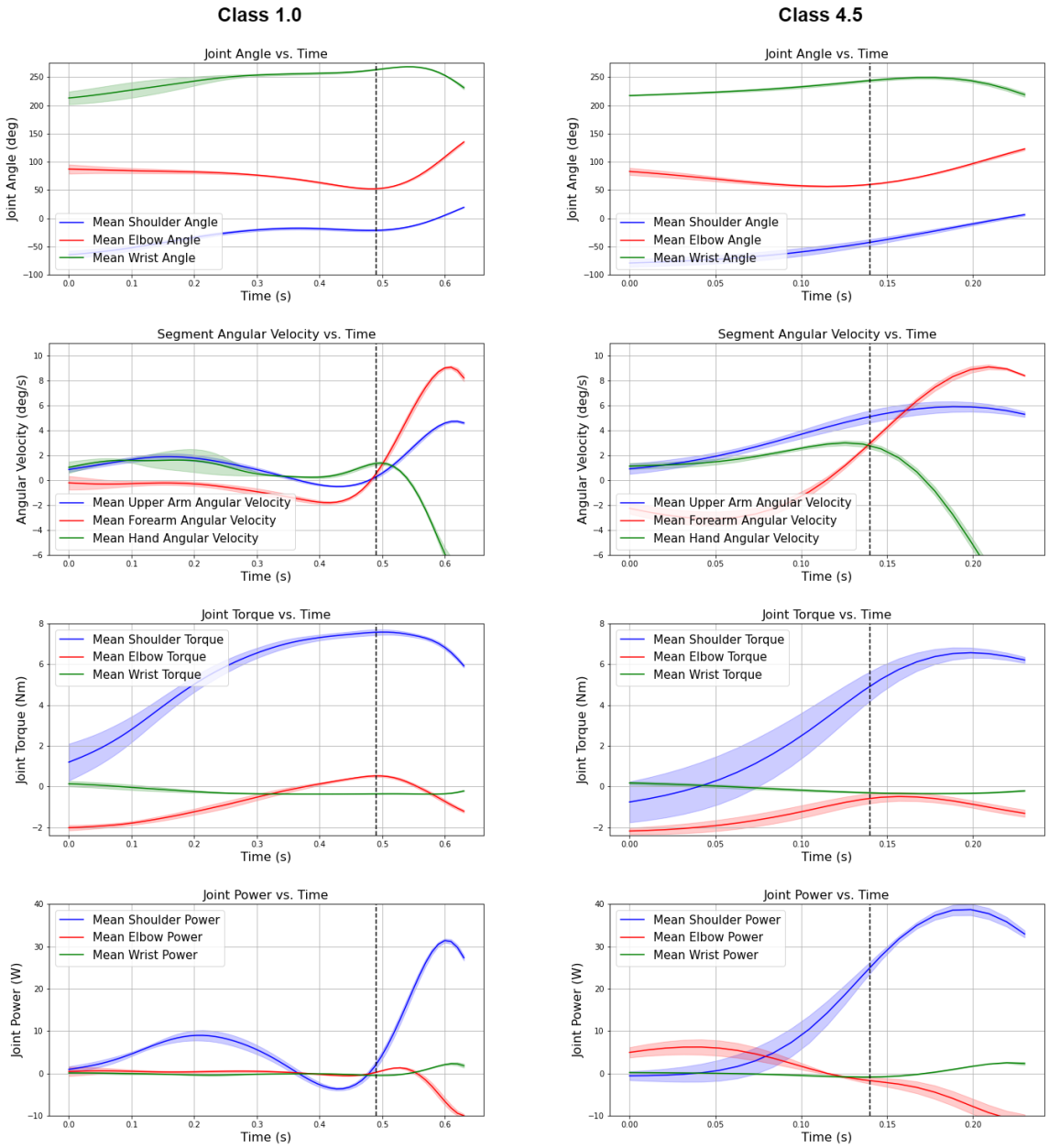


Figure 4.12: Differences between the class 1.0 (left column) and the class 4.5 shot (right column) shooting forms using pose estimation. (Shaded region represents ±1 std dev)

Limitations of Trunk Exclusion in Pose Estimation

The current implementation of pose estimation focuses exclusively on upper limb joints, omitting trunk kinematics from the analysis. This represents a significant limitation given that trunk movement and stability are fundamental to wheelchair basketball performance [198]. The IWBF classification system recognizes trunk control as the primary differentiator between player classes, with the "volume of action" - the extremes to which a player's trunk stability allows them to reach without overbalancing - serving as the key classification criterion [86].

The absence of trunk analysis in *ShotPose* has several implications:

1. **Incomplete Biomechanical Picture**: While the upper limb analysis reveals compensatory mechanisms, it cannot capture how class 4.5 athletes utilize trunk extension and rotation to generate additional force, nor how class 1.0 athletes stabilize their limited trunk function.

2. **Missed Coaching Opportunities**: Research by Marszalek et al. [133] demonstrated significant differences in trunk range of motion between wheelchair basketball classes. Without measuring these movements, coaches may miss opportunities to optimize shooting technique within each athlete's functional capacity.

3. **Classification Validation**: Santos et al. [172] showed that objective measures of trunk function correlate positively with IWBF classification. Including trunk metrics in *ShotPose* could provide quantitative support for classification decisions.

The technical challenges of implementing trunk pose estimation in wheelchair basketball are substantial. The wheelchair backrest occludes the lower spine, making hip joint detection particularly challenging. Additionally, the variety of trunk strapping configurations used by different athletes creates inconsistent visual patterns that complicate pose estimation algorithms. These challenges explain the current focus on upper limb analysis while highlighting important areas for future development.

Chapter 5

Discussion

To the knowledge of the author, this work pioneers the implementation of computer vision algorithms to optimize the free throw shooting mechanics in wheelchair basketball. The deployment of pose estimation on a mobile device allows coaches and athletes to leverage the benefits of markerless motion capture and integrate biomechanics to individualize training regimens in any environment with minimal setup time. This approach not only streamlines analysis for coaches but also democratizes access to advanced biomechanical tools, making such technology feasible for underfunded programs and amateur athletes. Furthermore, this framework can be adapted to other sports or physical rehabilitation programs, where individualized biomechanics are critical for performance improvement and injury prevention.

The experiments conducted in this study revealed key biomechanical differences in shooting techniques across player classifications and shooting styles. Pose estimation analysis demonstrated that *one-motion shots* were more efficient than *two-motion shots*, as they generated lower joint torques and maintained continuous momentum throughout the shooting motion. In contrast, *two-motion shots* exhibited a noticeable pause in upper arm movement, leading to increased torque demands, particularly at the elbow, to re-accelerate the ball. These differences highlight how variations in shooting mechanics can impact performance efficiency.

In classification-based comparisons, class 4.5 athletes consistently employed a one-motion shot, utilizing trunk movement to generate additional force. In contrast, the class 1.0 athlete, who had more severe functional impairments, exhibited a two-motion shot with significantly higher shoulder and elbow torques. The increased reliance on upper limb mechanics was necessary to compensate for the lack of trunk involvement, reinforcing the

role of classification-specific adaptations in shooting mechanics. These findings demonstrate the potential for pose estimation to uncover critical biomechanical trends, enabling coaches to optimize shooting techniques based on an athlete's classification and individual movement constraints.

Coaches can observe the motion of the shooting arm through each phase of the shot across athletes of different classifications. Dynamic information presented through the free throw shooting application allows a coach to experiment with various techniques and measure the efficiency of the form. By providing real-time, actionable feedback, the tool empowers athletes and coaches to make data-driven decisions in their training.

The primary limitation of the free throw mobile application is the pose estimation's reduced precision in joint detection, particularly when compared to the Vicon system. This discrepancy affects the accuracy of kinematic and dynamic calculations, diminishing the granularity of the analysis. Notably, RMSE for wrist angles exceeds 63°, primarily due to motion blur affecting distal joints during rapid movements. Additionally, differences in measurement methodologies between systems contribute to these variances. These limitations may also introduce bias if the training data does not adequately reflect the diversity of athletes in terms of mobility, skin tones, or environmental conditions. The performance of the application is further restricted by the quality of images and videos recorded by the device camera. Future iterations of this technology might incorporate pre-trained models tailored for specific populations, such as wheelchair athletes, and utilize higher frame rate cameras to address precision gaps. Enhancing pose estimation accuracy would improve its ability to capture the nuances of shooting mechanics, ultimately making it a more reliable tool for both research and coaching applications.

It is important to note that the validation against VICON, while providing a benchmark for comparison, has its own limitations due to the simplified marker protocol employed. Future validation studies could benefit from using ISB-compliant marker sets with multiple markers per segment, though this would increase complexity and potentially limit the practical coaching applications that are central to this thesis's objectives.

The absence of trunk analysis in the current version of *ShotPose* represents a significant limitation given the critical role of trunk function in wheelchair basketball shooting mechanics. As demonstrated in the case studies, class 4.5 athletes utilized trunk movement to generate additional force while maintaining lower joint torques, whereas the class 1.0 athlete compensated for limited trunk function through significantly higher shoulder and elbow torques. This compensation mechanism underscores the importance of trunk dynamics in understanding the complete biomechanical picture of wheelchair basketball shooting. Without trunk metrics, coaches may inadvertently focus exclusively on upper

limb mechanics, potentially missing opportunities to optimize shooting technique within each athlete's trunk functional capacity. Future iterations must address this limitation to provide comprehensive biomechanical analysis.

5.1 Comparison with State-of-the-Art Sports Pose Estimation Systems

To position *ShotPose* within the broader landscape of sports pose estimation, it is valuable to compare its approach and performance with recent advances in the field. Several notable systems have emerged that address similar challenges in sports biomechanics, each offering unique insights for wheelchair basketball applications.

Recent commercial systems like *PitchAI* [61] have demonstrated the feasibility of phone-based 3D pose estimation for sports analysis. Their baseball pitching analysis achieves joint angle RMSE values between 4.37° and 20.78° when validated against marker-based systems, with 80% of position measurements achieving sub-30mm error. While *PitchAI* operates in a different sporting context, their success in extracting 3D biomechanical data from single-camera smartphone footage validates the practical viability of mobile pose estimation for coaching applications. However, their system focuses on standing athletes without the unique occlusion challenges presented by wheelchairs.

The emergence of wheelchair-specific pose estimation research provides particularly relevant benchmarks for *ShotPose*. *WheelPose* [80] addresses the fundamental challenge that general-purpose pose estimation models underperform on wheelchair users by up to 27.54%. Their synthetic data generation approach and 23-keypoint model specifically designed for wheelchair users achieved 7.64% improvement in mean average precision compared to standard models. This research validates *ShotPose*'s specialized approach for wheelchair basketball, as it demonstrates that domain-specific models significantly outperform general-purpose systems for wheelchair athletes.

In the broader sports pose estimation landscape, *SportsPose* [83] provides a comprehensive 3D dataset with 176,000+ poses across multiple sports, achieving 34.5mm mean error in validation studies. Their multi-view approach and focus on dynamic athletic movements offer insights into handling fast-paced sports actions. However, their reliance on multi-camera setups contrasts with *ShotPose*'s single-camera approach, highlighting the trade-off between accuracy and practical deployment constraints.

Recent validation studies by Fukushima et al. [62] establish accuracy benchmarks for sports pose estimation, reporting mean absolute errors of 9.0±3.3° for sports movements

using OpenPose. These findings align closely with *ShotPose*'s shoulder and elbow angle errors (RMSE ¡20°), though the significantly higher wrist angle errors (¿63°) in *ShotPose* highlight the particular challenge of tracking distal joints during rapid basketball shooting motions.

The integration of advanced architectures in systems like *VoxelPose* [194] and its real-time variant *Faster VoxelPose* [213] demonstrates sophisticated approaches to 3D reconstruction from multiple viewpoints. While these systems achieve superior accuracy through multi-camera fusion, they require infrastructure that conflicts with *ShotPose*'s design goal of accessibility and minimal setup requirements.

Notably, *ShotPose* appears to be among the first systems specifically targeting wheelchair basketball biomechanics with real-time mobile deployment. While other systems have addressed wheelchair users (*WheelPose*) or basketball analysis separately, the combination of wheelchair-specific considerations, basketball shooting mechanics, and mobile deployment represents a unique contribution. The integration of inverse dynamics calculations for torque estimation further distinguishes *ShotPose* from purely kinematic analysis systems.

The comparison reveals that while *ShotPose*'s accuracy metrics are competitive with state-of-the-art systems for proximal joints, the high wrist angle errors represent a significant limitation compared to multi-camera or marker-based approaches. However, this trade-off enables practical deployment advantages that align with the thesis's goal of democratizing biomechanical analysis for wheelchair basketball coaching.

The next section suggests future directions to improve the functionality and performance of the free throw mobile application.

5.2 Future Research

5.2.1 Performance of Pose Estimation

The accuracy of a pose estimation network is significantly dependent on the quality and the quantity of the training data. The *COCO-Wholebody* dataset does not contain images of wheelchair basketball athletes. Augmenting the training data with images/videos of athletes (different skin tones, body types, clothing) shooting basketball free throws in various environments and lighting conditions would improve the accuracy and reliability of the pose estimation network within the mobile application. Annotation of the keypoints must be performed consistently to avoid introducing noise into the dataset. Ensuring

diversity in the training data could also reduce biases and increase the tool's generalizability across a wide range of use cases.

Recent advances in wheelchair-specific pose estimation, such as *WheelPose* [80], demonstrate that synthetic data generation can improve model performance by up to 27.54% for wheelchair users. Adopting similar data augmentation strategies specifically for wheelchair basketball scenarios could significantly enhance *ShotPose*'s accuracy. Additionally, the emergence of sports-specific datasets like *AthletePose3D* [214], which shows 69% reduction in error when models are fine-tuned on athletic data, reinforces the importance of domain-specific training data.

Since the motion of the human body occurs in 3D, the accuracy of joint detection can significantly increase by performing pose estimation in 3D. Current 3D pose estimation networks apply a scaling transformation to "lift" keypoints, detected by an underlying 2D pose estimator, from 2D to 3D. Thus, the accuracy of 3D pose reconstructions is directly influenced by the accuracy of 2D pose estimation and heavily biased by the scaling transformation [134]. However, advancements in smartphone technology, such as the integration of depth-sensing light detecting and ranging (LiDAR) sensors in devices like the iPhone 16, open new possibilities. Combining RGB video data with depth data could significantly enhance the precision of 3D pose estimation and enable real-time motion analysis in a broader range of environments [192].

5.2.2 Functionality of Free Throw Mobile App

Trunk Involvement

The biomechanical analysis of the free throw application must be extended to include trunk dynamics, addressing a fundamental limitation of the current system. The trunk plays a critical role in wheelchair basketball shooting mechanics, serving as both the primary source of momentum generation for higher-class athletes and a crucial stability platform for all players [198]. The IWBF classification system recognizes trunk control as the most important factor in player classification, with trunk movement assessment forming the cornerstone of functional evaluation [86].

Research has consistently demonstrated the biomechanical importance of trunk function in wheelchair basketball:

1. **Force Generation**: Gil et al. [67] found that players with higher classifications (greater trunk function) generated significantly greater maximal pass distances (r=0.67)

and medicine ball throw distances, directly attributable to their ability to utilize trunk rotation and extension.

2. **Shooting Efficiency**: Goosey-Tolfrey et al. [72] observed distinct biomechanical strategies between classification levels, with higher-class players utilizing trunk lean to achieve optimal release angles while maintaining lower joint velocities.

3. **Movement Coordination**: Santos et al. [172] quantified trunk strength differences between classes, showing progressive increases in peak trunk extensor and flexor torque that correlated with classification level.

Future iterations of *ShotPose* should prioritize trunk analysis through several approaches:

Shoulder Joint Center Trajectory Analysis: As an intermediate solution, tracking the motion of the shoulder joint center over time could serve as a proxy for trunk movement. The trajectory of this point reflects trunk lean, rotation, and elevation changes during the shooting motion. This approach avoids the technical challenges of hip detection while providing valuable information about trunk contribution to the shot.

Multi-Point Trunk Estimation: Advanced pose estimation models could be trained to detect multiple points along the visible portions of the trunk, including the sternum, upper spine markers, and lateral trunk boundaries. Recent work by Marszalek et al. [133] using Kinect sensors demonstrated reliable measurement of trunk movements in wheelchair basketball players, achieving ICC values above 0.75 for most movements.

Wheelchair-Specific Training Data: The effects of wheelchair occlusion on hip joint detection can be minimized through specialized training datasets. The emergence of wheelchair-specific pose estimation research, such as *WheelPose* [80], demonstrates that targeted training data can overcome occlusion challenges and improve keypoint detection accuracy by up to 27.54% for wheelchair users.

Integration with Classification Systems: Incorporating trunk metrics would allow *ShotPose* to provide quantitative data supporting classification decisions. The ability to objectively measure trunk range of motion, stability, and contribution to shooting mechanics could assist classifiers and provide athletes with clearer understanding of their functional capabilities.

The absence of trunk analysis in the current version of *ShotPose* represents not just a technical limitation but a missed opportunity to provide comprehensive biomechanical feedback. The inclusion of trunk metrics would direct appropriate coaching attention to this fundamental aspect of wheelchair basketball performance, potentially leading to more effective training strategies across all classification levels.

Research on occlusion-aware networks such as *CONet* [15] demonstrates that specialized architectures can handle occluded joints with significantly improved accuracy. Adapting these techniques for wheelchair-induced occlusions could address one of *ShotPose*'s key challenges in trunk analysis.

Action Segmentation Model

In the current version of the application, the user must scroll through the video to manually determine the time points representing the phases of the shot. This process can be automated using action segmentation. Action segmentation divides an untrimmed video by time and classifies each segmented part with a particular action (*i.e.* loading phase, transition phase, launch phase) using a computer vision model. Such automation would enable coaches to observe the duration of each phase of the free throw and how this impacts the efficiency of the shot, leading to more precise and efficient training interventions.

Shot-Learning

The application can also include a "shot-learning" feature. A shooter can record multiple successful shooting videos and apply pose estimation on each of them. The application would then average the joint positions across all of the shots to learn the shooting motion unique to that shooter. This learned motion could be superimposed on new shooting videos to highlight deviations in arm movement. Additionally, the shooter's learned motion could be compared to other shooters to identify key areas for improvement or optimization. Such features align with the contemporary trend of personalized coaching, offering athletes an AI-driven tool to refine their unique shooting styles over time.

5.2.3 Interdisciplinary Collaboration

Future research could also explore interdisciplinary collaborations to enhance the free throw application. For example, integrating psychological profiling with pose-based analysis could

uncover the interplay between mental states and biomechanical efficiency during high-pressure scenarios. Partnering with sports scientists, biomechanical engineers, and AI researchers could yield innovations in motion analysis and training methodologies.

The success of multi-disciplinary approaches in systems like *WheelPoser* [118], which combines IMU sensors with pose estimation for enhanced accuracy, suggests that sensor fusion could provide valuable enhancements to *ShotPose*. Integrating sparse IMU data could potentially address the high wrist angle errors while maintaining the system's accessibility.

5.2.4 Broader Vision

Ultimately, the integration of advanced pose estimation with ubiquitous mobile technologies promises to revolutionize athletic training, bridging the gap between elite and amateur sports while fostering inclusivity and accessibility. By continuously refining these tools, this work contributes to a future where data-driven coaching becomes a cornerstone of athletic performance and injury prevention.

References

[1] Athos apparel — smart clothing lab - smart clothing lab. https://smartclothinglab.com/brands/athos-apparel/. (Accessed on 1/18/2022).

[2] Automatic, real-time basketball stats and analytics — shottracker. https://shottracker.com/. (Accessed on 1/18/2022).

[3] Calvin upper-body detector v1.04. http://calvin-vision.net/bigstuff/calvin_upperbody_detector/. Accessed: 2020-04-15.

[4] Homecourt. https://www.homecourt.ai/. (Accessed on 10/18/2022).

[5] Models - machine learning - apple developer. https://developer.apple.com/machine-learning/models/. (Accessed on 10/25/2022).

[6] Tensorflow lite — ml for mobile and edge devices. https://www.tensorflow.org/lite. (Accessed on 10/25/2022).

[7] Proceedings ieee conference on advanced video and signal based surveillance. avss 2003. In *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, 2003.*, 2003.

[8] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[9] Ayman Ali, Ekkasit Pinyoanuntapong, Pu Wang, and Mohsen Dorodchi. A modular multi-stage lightweight graph transformer network for human pose and shape estimation from 2d human pose, 2023.

[10] Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, and Bernt Schiele. 2d human pose estimation: New benchmark and state of the art analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3686–3693, 2014.

[11] Mykhaylo Andriluka, Stefan Roth, and Bernt Schiele. Pictorial structures revisited: People detection and articulated pose estimation. pages 1014 – 1021, 07 2009.

[12] Bruno Artacho and Andreas Savakis. Unipose: Unified human pose estimation in single images and videos, 2020.

[13] Bruno Artacho and Andreas Savakis. Omnipose: A multi-scale framework for multi-person pose estimation, 2021.

[14] Marygrove College Athletics. In-depth: The science behind a basketball free throw, 2025.

[15] Jingrui Bai, Yu Ren, and Junliang Zhang. Conet: Crowd and occlusion-aware network for occluded human pose estimation. *Neural Networks*, 172:106099, 2024.

[16] Irina Barzykina. The physics of an optimal basketball free throw, 2017.

[17] Vasileios Belagiannis and Andrew Zisserman. Recurrent human pose estimation, 2017.

[18] Nicholai A Bernstein. On the construction of movements, 1947.

[19] Peter J. Brancazio. Physics of basketball. *American Journal of Physics*, 49(4):356–365, April 1981.

[20] Frank Brasile and Bradley Hedrick. The relationship of skills of elite wheelchair basketball competitors to the international functional classification system. *Therapeutic Recreation Journal*, 30(2), 1996.

[21] Brown, Colin. Predictive forward dynamic simulation of manual wheelchair propulsion. 2018.

[22] Adrian Bulat, Jean Kossaifi, Georgios Tzimiropoulos, and Maja Pantic. Toward fast and accurate human pose estimation via soft-gated skip connections, 2020.

[23] Chris Button, Malcolm MacLeod, Ross Sanders, and Simon Coleman. Examining movement variability in the basketball free-throw action at different skill levels. *Research Quarterly for Exercise and Sport*, 74(3):257–269, 2003.

[24] Wheelchair Basketball Canada. History of the sport, Mar 2016.

[25] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields, 2017.

[26] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with transformers, 2020.

[27] Joao Carreira, Pulkit Agrawal, Katerina Fragkiadaki, and Jitendra Malik. Human pose estimation with iterative error feedback, 2016.

[28] Haoming Chen, Runyang Feng, Sifan Wu, Hao Xu, Fengcheng Zhou, and Zhenguang Liu. 2d human pose estimation: A survey. *arXiv preprint arXiv:2204.07370*, 2022.

[29] Haoming Chen, Runyang Feng, Sifan Wu, Hao Xu, Fengcheng Zhou, and Zhenguang Liu. 2d human pose estimation: A survey, 2022.

[30] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L. Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, 2017.

[31] Lulu Chen, Hong Wei, and James Ferryman. A survey of human motion analysis using depth imagery. *Pattern Recognition Letters*, 34(15):1995–2006, 2013.

[32] Mark Chen, Alec Radford, Rewon Child, Jeffrey Wu, Heewoo Jun, David Luan, and Ilya Sutskever. Generative pretraining from pixels. In *International Conference on Machine Learning*, pages 1691–1703. PMLR, 2020.

[33] Xianjie Chen and Alan Yuille. Articulated pose estimation by a graphical model with image dependent pairwise relations, 2014.

[34] Yilun Chen, Zhicheng Wang, Yuxiang Peng, Zhiqiang Zhang, Gang Yu, and Jian Sun. Cascaded pyramid network for multi-person pose estimation, 2018.

[35] Yu Chen, Chunhua Shen, Xiu-Shen Wei, Lingqiao Liu, and Jian Yang. Adversarial posenet: A structure-aware convolutional network for human pose estimation, 2017.

[36] Yucheng Chen, Yingli Tian, and Mingyi He. Monocular human pose estimation: A survey of deep learning-based methods. *Computer Vision and Image Understanding*, 192:102897, mar 2020.

[37] Bowen Cheng, Bin Xiao, Jingdong Wang, Honghui Shi, Thomas S. Huang, and Lei Zhang. Higherhrnet: Scale-aware representation learning for bottom-up human pose estimation, 2020.

[38] Chia-Jung Chou, Jui-Ting Chien, and Hwann-Tzong Chen. Self adversarial training for human pose estimation, 2017.

[39] Xiao Chu, Wanli Ouyang, Hongsheng Li, and Xiaogang Wang. Structured feature learning for pose estimation, 2016.

[40] Xiao Chu, Wei Yang, Wanli Ouyang, Cheng Ma, Alan L. Yuille, and Xiaogang Wang. Multi-context attention for human pose estimation, 2017.

[41] Steffi L Colyer, Murray Evans, Darren P Cosker, and Aki IT Salo. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports Medicine - Open*, 4(1):1–15, 2018.

[42] Steffi L Colyer, Murray Evans, Darren P Cosker, and Aki IT Salo. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports medicine-open*, 4(1):1–15, 2018.

[43] MMPose Contributors. Openmmlab pose estimation toolbox and benchmark. https://github.com/open-mmlab/mmpose, 2020.

[44] B Courbariaux. The basics of classification in wheelchair basketball, 2022.

[45] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893 vol. 1, 2005.

[46] Qi Dang, Jianqin Yin, Bin Wang, and Wenqing Zheng. Deep learning based 2d human pose estimation: A survey. *Tsinghua Science and Technology*, 24(6):663–676, 2019.

[47] Matthias Dantone, Juergen Gall, Christian Leistner, and Luc Van Gool. Human pose estimation using body parts dependent joint regressors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3041–3048, 2013.

[48] Paolo De Leva. Adjustments to zatsiorsky-seluyanov's segment inertia parameters. *Journal of biomechanics*, 29(9):1223–1230, 1996.

[49] Bappaditya Debnath, Mary O'Brien, Motonori Yamaguchi, and Ardhendu Behera. Adapting mobilenets for mobile based upper body pose estimation. In *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 1–6. IEEE, 2018.

[50] Demirdjian, Ko, and Darrell. Constraining human body tracking. In *Proceedings Ninth IEEE International Conference on Computer Vision*, pages 1071–1078 vol.2, 2003.

[51] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[52] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale, 2020.

[53] Douglas Brown. Physlets tracker: Video analysis and modeling tool.

[54] Bruce Elliott. A kinematic and kinetic analysis of the female two point and three point jump shots in basketball. *Australian Journal of Science and Medicine in Sport*, 24:111–111, 1992.

[55] Bruce Elliott, University of Western Australia., and National Sports Research Centre (Australia). *The jump shot : a comparison of male and female shooting techniques / by Bruce Elliott*. National Sports Research Centre, [199-?].

[56] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Neural architecture search: A survey, 2019.

[57] Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, and Cewu Lu. Rmpe: Regional multi-person pose estimation, 2018.

[58] Pedro F Felzenszwalb, Ross B Girshick, David McAllester, and Deva Ramanan. Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*, 32(9):1627–1645, 2009.

[59] Pedro F Felzenszwalb and Daniel P Huttenlocher. Pictorial structures for object recognition. *International journal of computer vision*, 61(1):55–79, 2005.

[60] P.F. Felzenszwalb and D.P. Huttenlocher. Efficient matching of pictorial structures. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, volume 2, pages 66–73 vol.2, 2000.

[61] Glenn S Fleisig, Jonathan S Slowik, Dan Wassom, Yuki Yanagita, Justin Bishop, and Alek Diffendaffer. Validation of pitchai markerless motion capture using marker-based 3d motion capture. *Sports Biomechanics*, pages 1–14, 2022.

[62] Aiden Fukushima and Hiroki Nakata. The potential of human pose estimation for motion capture in sports: a validation study. *Sports Engineering*, 27(1):1–12, 2024.

[63] Michael Fürst, Shriya T. P. Gupta, René Schuster, Oliver Wasenmüller, and Didier Stricker. Hperl: 3d human pose estimation from rgb and lidar, 2020.

[64] Joerg M. Gablonsky and Andrew S. I. D. Lang. Modeling basketball free throws. *SIAM Review*, 47(4):775–798, January 2005.

[65] Walter H Gage, David A Winter, James S Frank, and Allan L Adkin. Synchronization of video kinematic and analog biomechanical data using the motion analysis system. In *ISBS-Conference Proceedings Archive*, 1991.

[66] D. Gerónimo, A. M. López, A. D. Sappa, and T. Graf. Survey of pedestrian detection for advanced driver assistance systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(7):1239–1258, 2010.

[67] Susana M Gil, Javier Yanci, Magdalena Otero, Jurgi Olasagasti, Aitor Badiola, Iraia Bidaurrazaga-Letona, Aitor Iturricastillo, and Cristina Granados. The functional classification and field test performance in wheelchair basketball players. *Journal of Human Kinetics*, 46:219–230, 2015.

[68] Georgia Gkioxari, Bharath Hariharan, Ross Girshick, and Jitendra Malik. R-cnns for pose estimation and action detection, 2014.

[69] Georgia Gkioxari, Alexander Toshev, and Navdeep Jaitly. Chained predictions using convolutional neural networks, 2016.

[70] Wenjuan Gong, Xuena Zhang, Jordi Gonzàlez, Andrews Sobral, Thierry Bouwmans, Changhe Tu, and El-hadi Zahzah. Human pose estimation from monocular images: A comprehensive survey. *Sensors*, 16(12):1966, 2016.

[71] Victoria Goosey-Tolfrey, Daniel Butterworth, and Calvin Morriss. Free throw shooting technique of male wheelchair basketball players. *Adapted Physical Activity Quarterly*, 19(2):238–250, 2002.

[72] Victoria Goosey-Tolfrey, David Butterworth, and Chris Morriss. Free throw shooting technique of male wheelchair basketball players. *Adapted Physical Activity Quarterly*, 19(2):238–250, 2002.

[73] Michel Goossens, Frank Mittelbach, and Alexander Samarin. *The LaTeX Companion*. Addison-Wesley, Reading, Massachusetts, 1994.

[74] George E Gorton III, David A Hebert, and Mary E Gannotti. Assessment of the kinematic variability among 12 motion analysis laboratories. *Gait & Posture*, 29(3):398–402, 2009.

[75] Gordon R Hamilton and Christoph Reinschmidt. Optimal trajectory for the basketball free throw. *Journal of Sports Sciences*, 15(5):491–504, 1997.

[76] Herbert Hatze. The fundamental problem of myoskeletal inverse dynamics and its implications. *Journal of biomechanics*, 35(1):109–115, 2002.

[77] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn, 2018.

[78] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.

[79] Thomas Hellsten, Jonny Karlsson, Muhammed Shamsuzzaman, and Göran Pulkkis. The potential of computer vision-based marker-less human motion analysis for rehabilitation. *Rehabilitation Process and Outcome*, 10:11795727211022330, 2021.

[80] William Huang, Rijin Sam, Zhi Gharaee, Koji Kajita, Peter Liang, Aishat Aloba, James Zou, and Victor Dibia. Wheelpose: Data synthesis techniques to improve pose estimation performance on wheelchair users. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2024.

[81] Loc N Huynh, Youngki Lee, and Rajesh Krishna Balan. Deepmon: Mobile gpu-based deep learning framework for continuous vision applications. In *Proceedings*

of the 15th Annual International Conference on Mobile Systems, Applications, and Services, pages 82–95, 2017.

[82] Snowden S. Ingram B.

[83] Christian Keilstrup Ingwersen, Rasmus Aagaard Jørgensen, Morten Rieger Mikkelstrup, Janus Nørtoft Jensen, Morten Rahr Hannemose, and Anders Bjorholm Dahl. Sportspose: A dynamic 3d sports pose dataset. *arXiv preprint arXiv:2304.01865*, 2023.

[84] Eldar Insafutdinov, Mykhaylo Andriluka, Leonid Pishchulin, Siyu Tang, Evgeny Levinkov, Bjoern Andres, and Bernt Schiele. Arttrack: Articulated multi-person tracking in the wild, 2017.

[85] Eldar Insafutdinov, Leonid Pishchulin, Bjoern Andres, Mykhaylo Andriluka, and Bernt Schiele. Deepercut: A deeper, stronger, and faster multi-person pose estimation model, 2016.

[86] International Wheelchair Basketball Federation. *Official Player Classification Manual*. IWBF, 2014.

[87] International Wheelchair Basketball Federation. Classification, 2025.

[88] Umar Iqbal and Juergen Gall. Multi-person pose estimation with local joint-to-person associations, 2016.

[89] International Wheelchair Basketball Federation (IWBF). Wheelchair basketball classification explained, 2023.

[90] Arjun Jain, Jonathan Tompson, Mykhaylo Andriluka, Graham W. Taylor, and Christoph Bregler. Learning human pose estimation features with convolutional networks, 2014.

[91] Arjun Jain, Jonathan Tompson, Yann LeCun, and Christoph Bregler. Modeep: A deep learning framework using motion features for human pose estimation, 2014.

[92] Xiaofei Ji and Honghai Liu. Advances in view-invariant human motion analysis: A review. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(1):13–24, 2009.

[93] Sheng Jin, Wentao Liu, Enze Xie, Wenhai Wang, Chen Qian, Wanli Ouyang, and Ping Luo. Differentiable hierarchical graph grouping for multi-person pose estimation, 2020.

[94] Sheng Jin, Xujie Ma, Zhipeng Han, Y. Wu, W. Yang, Wentao Liu, Chen Qian, and Wanli Ouyang. Towards multi-person pose tracking : Bottom-up and top-down methods. 2017.

[95] Sheng Jin, Lumin Xu, Jin Xu, Can Wang, Wentao Liu, Chen Qian, Wanli Ouyang, and Ping Luo. Whole-body human pose estimation in the wild. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020.

[96] Sam Johnson and Mark Everingham. Clustered pose and nonlinear appearance models for human pose estimation. In *Proceedings of the British Machine Vision Conference*, pages 12.1–12.11. BMVA Press, 2010. doi:10.5244/C.24.12.

[97] S. X. Ju, M. J. Black, and Y. Yacoob. Cardboard people: a parameterized model of articulated image motion. In *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, pages 38–44, 1996.

[98] Robert M Kanko, Elise K Laende, Elysia M Davis, W Scott Selbie, and Kevin J Deluzio. Comparison of kinematics and joint moments calculations for lower limbs during gait using markerless and marker-based motion capture. *Frontiers in Bioengineering and Biotechnology*, 12:1280363, 2023.

[99] Lipeng Ke, Ming-Ching Chang, Honggang Qi, and Siwei Lyu. Multi-scale structure-aware network for human pose estimation, 2018.

[100] Alex Kendall, Matthew Grimes, and Roberto Cipolla. Posenet: A convolutional network for real-time 6-dof camera relocalization, 2015.

[101] Duane Knudson. Biomechanics of the basketball jump shot—six key teaching points. *Journal of Physical Education, Recreation & Dance*, 64(2):67–73, 1993.

[102] Duane Knudson. Biomechanics of the basketball jump shot—six key teaching points. *Journal of Physical Education, Recreation & Dance*, 64(2):67–73, 1993.

[103] Donald Knuth. *The TEXbook*. Addison-Wesley, Reading, Massachusetts, 1986.

[104] Muhammed Kocabas, Salih Karagoz, and Emre Akbas. Multiposenet: Fast multi-person pose estimation using pose residual network, 2018.

[105] Vicki Komisar, William E McIlroy, and Carolyn A Duncan. A novel method for synchronizing motion capture with other data sources for millisecond-level precision. *Gait & Posture*, 51:125–131, 2017.

[106] Jean Kossaifi, Adrian Bulat, Georgios Tzimiropoulos, and Maja Pantic. T-net: Parametrizing fully convolutional nets with a single high-order tensor, 2019.

[107] Jerry Krause, Don Meyer, and Jerry Meyer. *Basketball skills and drills*. Human Kinetics, 2008.

[108] Sven Kreiss, Lorenzo Bertoni, and Alexandre Alahi. Pifpaf: Composite fields for human pose estimation, 2019.

[109] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[110] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Commun. ACM*, 60(6):84–90, may 2017.

[111] Leslie Lamport. *LaTeX — A Document Preparation System*. Addison-Wesley, Reading, Massachusetts, second edition, 1994.

[112] Nicholas D Lane, Sourav Bhattacharya, Petko Georgiev, Claudio Forlivesi, Lei Jiao, Lorena Qendro, and Fahim Kawsar. Deepx: A software accelerator for low-power deep learning inference on mobile devices. In *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 1–12. IEEE, 2016.

[113] Alberto Leardini, Lorenzo Chiari, Ugo Della Croce, and Aurelio Cappozzo. Human movement analysis using stereophotogrammetry. *Gait &amp Posture*, 21(2):212–225, February 2005.

[114] Alberto Leardini, Lorenzo Chiari, Ugo Della Croce, and Aurelio Cappozzo. Human movement analysis using stereophotogrammetry: Part 3. soft tissue artifact assessment and compensation. *Gait & posture*, 21(2):212–225, 2005.

[115] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[116] Bo Li, Huahui Chen, Yucheng Chen, Yuchao Dai, and Mingyi He. Skeleton boxes: Solving skeleton based action detection with a single deep convolutional neural network, 2017.

[117] Sijin Li, Zhi-Qiang Liu, and Antoni B. Chan. Heterogeneous multi-task learning for human pose estimation with deep convolutional neural network, 2014.

[118] Yunzhi Li, Vimal Mollyn, Kuang Yuan Li, Yuntao Kong, Mayank Goel, Shwetak Patel, Chris Harrison, and Karan Ahuja. Wheelposer: Sparse-imu based body pose estimation for wheelchair users. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–16, 2024.

[119] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection, 2017.

[120] Hui Liu. Chapter 3 - rail transit collaborative robot systems. In Hui Liu, editor, *Robot Systems for Rail Transit Applications*, pages 89–141. Elsevier, 2020.

[121] Zhao Liu, Jianke Zhu, Jiajun Bu, and Chun Chen. A survey of human pose estimation: the body parts parsing based methods. *Journal of Visual Communication and Image Representation*, 32:10–19, 2015.

[122] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation, 2015.

[123] D.G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 2, pages 1150–1157 vol.2, 1999.

[124] Yue Luo, Jimmy Ren, Zhouxia Wang, Wenxiu Sun, Jinshan Pan, Jianbo Liu, Jiahao Pang, and Liang Lin. Lstm pose machines, 2018.

[125] Diogo C. Luvizon, David Picard, and Hedi Tabia. 2d/3d pose estimation and action recognition using multitask deep learning, 2018.

[126] Diogo C. Luvizon, Hedi Tabia, and David Picard. Human pose regression by combining indirect part detection and contextual information, 2017.

[127] L Malone, P Gervais, and R Steadward. Parameters of ball release in wheelchair basketball free throw shooting. In *ISBS-Conference Proceedings Archive*, 1999.

[128] LA Malone, PL Gervais, JP Baudin, and R Steadward. Klnematics of free throw shooting by class 1.0 wheelchair basketball players. In *ISBS-Conference Proceedings Archive*, 1995.

[129] Laurie A Malone, Pierre L Gervais, and Robert D Steadward. Shooting mechanics related to player classification and free throw success in wheelchair basketball. *Journal of Rehabilitation Research and Development*, 39(6):701–710, 2002.

[130] Laurie A Malone, Pierre L Gervais, and Robert D Steadward. Shooting mechanics related to player classification and free throw success in wheelchair basketball. *Journal of Rehabilitation Research and Development*, 39(6):701–710, 2002.

[131] Laurie A Malone, A Brian Nielsen, and Robert D Steadward. Expanding the dichotomous outcome in wheelchair basketball shooting of elite male players. *Adapted Physical Activity Quarterly*, 17(4):437–449, 2000.

[132] E. Marinoiu, D. Papava, and C. Sminchisescu. Pictorial human spaces: How well do humans perceive a 3d articulated pose? In *2013 IEEE International Conference on Computer Vision*, pages 1289–1296, 2013.

[133] Jolanta Marszałek and Bartosz Molik. Reliability of measurement of active trunk movement in wheelchair basketball players. *PLoS One*, 14(11):e0225515, 2019.

[134] Julieta Martinez, Rayat Hossain, Javier Romero, and James J Little. A simple yet effective baseline for 3d human pose estimation. In *Proceedings of the IEEE international conference on computer vision*, pages 2640–2649, 2017.

[135] V Maynard, A.M.O Bakheit, J Oldham, and J Freeman. Intra-rater and inter-rater reliability of gait measurements with CODA mpx30 motion analysis system. *Gait &amp Posture*, 17(1):59–67, February 2003.

[136] William McNally, Kanav Vats, Alexander Wong, and John McPhee. EvoPose2d: Pushing the boundaries of 2d human pose estimation using accelerated neuroevolution with weight transfer. *IEEE Access*, 9:139403–139414, 2021.

[137] William McNally, Kanav Vats, Alexander Wong, and John McPhee. Rethinking keypoint representations: Modeling keypoints and poses as objects for multi-person human pose estimation, 2021.

[138] MMPose Contributors. MMPose Documentation: Wholebody 2D Keypoint Models. https://mmpose.readthedocs.io/en/latest/model_zoo/wholebody_2d_keypoint.html, 2024.

[139] Thomas B Moeslund and Erik Granum. A survey of computer vision-based human motion capture. *Computer vision and image understanding*, 81(3):231–268, 2001.

[140] Thomas B Moeslund, Adrian Hilton, and Volker Krüger. A survey of advances in vision-based human motion capture and analysis. *Computer vision and image understanding*, 104(2-3):90–126, 2006.

[141] Thomas B Moeslund, Adrian Hilton, Volker Krüger, and Leonid Sigal. *Visual analysis of humans*. Springer, 2011.

[142] Pavlo Molchanov, Arun Mallya, Stephen Tyree, Iuri Frosio, and Jan Kautz. Importance estimation for neural network pruning, 2019.

[143] Gyeongsik Moon, Ju Yong Chang, and Kyoung Mu Lee. Posefix: Model-agnostic general human pose refinement network, 2019.

[144] Melissa M.B. Morrow, Bethany Lowndes, Emma Fortune, Kenton R. Kaufman, and M. Susan Hallbeck. Validation of inertial measurement units for upper body kinematics. *Journal of Applied Biomechanics*, 33(3):227–232, June 2017.

[145] T. L. Munea, Y. Z. Jembre, H. T. Weldegebriel, L. Chen, C. Huang, and C. Yang. The progress of human pose estimation: A survey and taxonomy of models applied in 2d human pose estimation. *IEEE Access*, 8:133330–133348, 2020.

[146] Tewodros Legesse Munea, Zelalem Yalew, Halefom Tekle, Longbiao Chen, Chenxi Huang, and Chenhui Yang. The progress of human pose estimation: A survey and taxonomy of models applied in 2d human pose estimation. *IEEE Access*, PP:1–1, 07 2020.

[147] Ryan Needham, Roozbeh Naemi, Joseph Hamill, and Nachiappan Chockalingam. Validity and reliability of trunk and lower-limb kinematics during squatting, hopping, jumping and side-stepping using opencap markerless motion capture application. *Journal of Sports Sciences*, 42(3):234–248, 2024.

[148] Alejandro Newell, Zhiao Huang, and Jia Deng. Associative embedding: End-to-end learning for joint detection and grouping, 2017.

[149] Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked hourglass networks for human pose estimation, 2016.

[150] Aiden Nibali, Zhen He, Stuart Morgan, and Luke Prendergast. Numerical coordinate regression with convolutional neural networks, 2018.

[151] Guanghan Ning, Zhi Zhang, and Zhihai He. Knowledge-guided deep fractal neural networks for human pose estimation, 2017.

[152] V. H. Okazaki and A. L. Rodacki. Increased distance of shooting on basketball jump shot. *J Sports Sci Med*, 11(2):231–237, 2012.

[153] Hiroki Okubo and Mont Hubbard. Kinematics of arm joint motions in basketball shooting. *Procedia Engineering*, 112:443–448, 2015.

[154] Hiroki Okubo and Mont Hubbard. Analysis of arm joint torques at ball-release for set and jump shots in basketball. In *The 13th Conference of the International Sports Engineering Association*. MDPI, June 2020.

[155] Keiron O'Shea and Ryan Nash. An introduction to convolutional neural networks, 2015.

[156] Tarik Ozmen, Bekir Yuktasir, Necmiye Un Yildirim, Birol Yalcin, and Mark ET Willems. Trunk function correlates positively with wheelchair basketball player classification. *American Journal of Physical Medicine & Rehabilitation*, 96(8):579–583, 2017.

[157] George Papandreou, Tyler Zhu, Liang-Chieh Chen, Spyros Gidaris, Jonathan Tompson, and Kevin Murphy. Personlab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model, 2018.

[158] George Papandreou, Tyler Zhu, Nori Kanazawa, Alexander Toshev, Jonathan Tompson, Chris Bregler, and Kevin Murphy. Towards accurate multi-person pose estimation in the wild, 2017.

[159] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[160] Georgios Pavlakos, Luyang Zhu, Xiaowei Zhou, and Kostas Daniilidis. Learning to estimate 3d human pose and shape from a single color image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 459–468, 2018.

[161] Xi Peng, Zhiqiang Tang, Fei Yang, Rogerio Feris, and Dimitris Metaxas. Jointly optimize data augmentation and network training: Adversarial data augmentation in human pose estimation, 2018.

[162] Xavier Perez-Sala, Sergio Escalera, Cecilio Angulo, and Jordi Gonzalez. A survey on model based approaches for 2d and 3d visual human pose recovery. *Sensors*, 14(3):4189–4210, 2014.

[163] Tomas Pfister, James Charles, and Andrew Zisserman. Flowing convnets for human pose estimation in videos, 2015.

[164] Leonid Pishchulin, Eldar Insafutdinov, Siyu Tang, Bjoern Andres, Mykhaylo Andriluka, Peter V Gehler, and Bernt Schiele. Deepcut: Joint subset partition and labeling for multi person pose estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4929–4937, 2016.

[165] Ronald Poppe. Vision-based human motion analysis: An overview. *Computer vision and image understanding*, 108(1-2):4–18, 2007.

[166] Nikola Prvulovic, Miljan Hadzovic, and Ana Lilic. A biomechanical analysis of the free throw shooting technique in wheelchair basketball: A pilot study.

[167] Lingteng Qiu, Xuanye Zhang, Yanran Li, Guanbin Li, Xiaojun Wu, Zixiang Xiong, Xiaoguang Han, and Shuguang Cui. Peeking into occluded joints: A novel framework for crowd pose estimation, 2020.

[168] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks, 2016.

[169] Xavier Robert-Lachaine, Hakim Mecheri, Christian Larue, and André Plamondon. Validation of a low-cost inertial motion capture system for whole-body motion analysis. *Journal of Biomechanics*, 99:109449, 2017.

[170] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[171] Sebastian Ruder. An overview of multi-task learning in deep neural networks, 2017.

[172] Ali Saltan and Handan Ankarali. The role of trunk stabilization in functional-classification levels in wheelchair basketball. *Journal of Sport Rehabilitation*, 26(4):287–293, 2017.

[173] Brianne N Schwark, Sasho J Mackenzie, and Eric J Sprigings. Optimizing the release conditions for a free throw in wheelchair basketball. *Journal of Applied Biomechanics*, 20(2), 2004.

[174] Michael H Schwartz, Joyce P Trost, and Roy A Wervey. Measurement and management of errors in quantitative gait data. *Gait & Posture*, 20(2):196–203, 2004.

[175] Anat Shkedy Rabani, Sarai Mizrachi, Gregory S Sawicki, and Raziel Riemer. Parametric equations to study and predict lower-limb joint kinematics and kinetics during human walking and slow running on slopes. *PloS one*, 17(8):e0269061, 2022.

[176] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.

[177] Ke Song, Christopher J Robinson, Gabriel Anand Orozco, Derrick M Dalton, Junfing Ko, Jonathan Mao, and Lindsay V Slater. Validation of opencap: A low-cost markerless motion capture system for lower-extremity kinematics during return-to-sport tasks. *Journal of Biomechanics*, 153:111600, 2024.

[178] Lucas Stoffl, Maxime Vidal, and Alexander Mathis. End-to-end trainable multi-instance pose estimation with transformers, 2021.

[179] Kai Su, Dongdong Yu, Zhenqi Xu, Xin Geng, and Changhu Wang. Multi-person pose estimation with enhanced channel-wise and spatial information, 2019.

[180] Guanxiong Sun, Chengqin Ye, and Kuanquan Wang. Focus on what's important: Self-attention model for human pose estimation, 2019.

[181] Ke Sun, Bin Xiao, Dong Liu, and Jingdong Wang. Deep high-resolution representation learning for human pose estimation, 2019.

[182] Xiao Sun, Jiaxiang Shang, Shuang Liang, and Yichen Wei. Compositional human pose regression, 2017.

[183] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions, 2014.

[184] Wei Tang and Ying Wu. Does learning specific features for related parts help human pose estimation? In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1107–1116, 2019.

[185] Wei Tang, Pei Yu, and Ying Wu. Deeply learned compositional models for human pose estimation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 190–206, 2018.

[186] Zhiqiang Tang, Xi Peng, Shijie Geng, Lingfei Wu, Shaoting Zhang, and Dimitris Metaxas. Quantized densely connected u-nets for efficient landmark localization, 2018.

[187] K. Tashiro, T. Kawamura, Y. Sei, H. Nakagawa, Y. Tahara, and A. Ohsuga. Refinement of ontology-constrained human pose classification. In *2014 IEEE International Conference on Semantic Computing*, pages 60–67, 2014.

[188] Katherine Thomas Thomas. The development of sport expertise: From leeds to MVP legend. *Quest*, 46(2):199–210, May 1994.

[189] Jonathan Tompson, Ross Goroshin, Arjun Jain, Yann LeCun, and Christopher Bregler. Efficient object localization using convolutional networks, 2015.

[190] Jonathan Tompson, Arjun Jain, Yann LeCun, and Christoph Bregler. Joint training of a convolutional network and a graphical model for human pose estimation, 2014.

[191] Alexander Toshev and Christian Szegedy. Deeppose: Human pose estimation via deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1653–1660, 2014.

[192] Alexander Toshev and Christian Szegedy. Human pose estimation via deep neural networks. *arXiv preprint arXiv:2010.08221*, 2020.

[193] Chau M. Tran and Larry M. Silverberg. Optimal release conditions for the free throw in men's basketball. *Journal of Sports Sciences*, 26(11):1147–1155, September 2008.

[194] Hanyue Tu, Chunyu Wang, and Wenjun Zeng. Voxelpose: Towards multi-camera 3d human pose estimation in wild environment. In *European Conference on Computer Vision*, pages 197–212. Springer, 2020.

[195] Scott D Uhlrich, Antoine Falisse, Lukasz Kidzinski, Julie Muccini, Michael Ko, Ajay S Chaudhari, Jennifer L Hicks, and Scott L Delp. Opencap: 3d human movement dynamics from smartphone videos. *PLoS Computational Biology*, 19(10):e1011462, 2023.

[196] Rafi Umer, Andreas Doering, Bastian Leibe, and Juergen Gall. Self-supervised keypoint correspondences for multi-person pose estimation and tracking in videos. *arXiv preprint arXiv:2004.12652*, 2020.

[197] Mehmet Uygur, Ayhan Goktepe, Emre Ak, Hakan Karabörk, and Feza Korkusuz. The effect of fatigue on the kinematics of free throw shooting in basketball. *Journal of human kinetics*, 24(2010):51–56, 2010.

[198] Yves C Vanlandewijck, Joeri Verellen, and Sean Tweedy. The relationship between functional potential and field performance in elite female wheelchair basketball players. *Journal of Sports Sciences*, 29(8):843–854, 2011.

[199] Fang Wang and Yi Li. Beyond physical connections: Tree models in human pose estimation, 2013.

[200] Pichao Wang, Wanqing Li, Philip Ogunbona, Jun Wan, and Sergio Escalera. Rgb-d-based human motion recognition with deep learning: A survey, 2018.

[201] Shih-En Wei, Varun Ramakrishna, Takeo Kanade, and Yaser Sheikh. Convolutional pose machines, 2016.

[202] Edward Joseph Williams. *Conquering the charity stripe: A literature review on the biomechanics of free throw shooting*. PhD thesis, 2021.

[203] David A Winter. *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[204] Ge Wu, Frans CT Van der Helm, HEJ DirkJan Veeger, Mohsen Makhsous, Peter Van Roy, Carolyn Anglin, Jochem Nagels, Andrew R Karduna, Kevin McQuade, Xuguang Wang, et al. Isb recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—part ii: shoulder, elbow, wrist and hand. *Journal of Biomechanics*, 38(5):981–992, 2005.

[205] Bin Xiao, Haiping Wu, and Yichen Wei. Simple baselines for human pose estimation and tracking, 2018.

[206] Lechao Xiao, Yasaman Bahri, Jascha Sohl-Dickstein, Samuel S. Schoenholz, and Jeffrey Pennington. Dynamical isometry and a mean field theory of cnns: How to train 10,000-layer vanilla convolutional neural networks, 2018.

[207] Lumin Xu, Yingda Guan, Sheng Jin, Wentao Liu, Chen Qian, Ping Luo, Wanli Ouyang, and Xiaogang Wang. Vipnas: Efficient video pose estimation via neural architecture search, 2021.

[208] Lumin Xu, Sheng Jin, Wentao Liu, Chen Qian, Wanli Ouyang, Ping Luo, and Xiaogang Wang. Zoomnas: Searching for whole-body human pose estimation in the wild. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[209] Sen Yang, Zhibin Quan, Mu Nie, and Wankou Yang. Transpose: Towards explainable human pose estimation by transformer, 2020.

[210] Wei Yang, Shuang Li, Wanli Ouyang, Hongsheng Li, and Xiaogang Wang. Learning feature pyramids for human pose estimation, 2017.

[211] Wei Yang, Wanli Ouyang, Hongsheng Li, and Xiaogang Wang. End-to-end learning of deformable mixture of parts and deep convolutional neural networks for human pose estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[212] Yi Yang and Deva Ramanan. Articulated pose estimation with flexible mixtures-of-parts. In *CVPR 2011*, pages 1385–1392. IEEE, 2011.

[213] Hang Ye, Wentao Zhu, Chunyu Wang, Rujie Wu, and Yizhou Wang. Faster voxelpose: Real-time 3d human pose estimation by orthographic projection. *arXiv preprint arXiv:2207.10955*, 2022.

[214] Jiaqi Yeung, Kaixiang Wang, Lan Li, Fan Yang, Xuefeng Gao, Ye Ji, and Kaiyu Shi. Athletepose3d: A benchmark dataset for 3d human pose estimation and kinematic validation in athletic movements. *arXiv preprint arXiv:2503.07499*, 2025.

[215] La Bar R.H. Dangelmaier B.S. Yilla, A.B. *Setting up a wheelchair for basketball*, volume 3. Sports'N Spokes, 1998.

[216] Feng Zhang, Xiatian Zhu, Hanbin Dai, Mao Ye, and Ce Zhu. Distribution-aware coordinate representation for human pose estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[217] Feng Zhang, Xiatian Zhu, and Mao Ye. Fast human pose estimation, 2019.

[218] Hong Zhang, Hao Ouyang, Shu Liu, Xiaojuan Qi, Xiaoyong Shen, Ruigang Yang, and Jiaya Jia. Human pose estimation with spatial contextual information, 2019.

[219] Yuexi Zhang, Yin Wang, Octavia Camps, and Mario Sznaier. Key frame proposal network for efficient pose estimation in videos, 2020.

[220] Ce Zheng, Wenhan Wu, Taojiannan Yang, Sijie Zhu, Chen Chen, Ruixu Liu, Ju Shen, Nasser Kehtarnavaz, and Mubarak Shah. Deep learning-based human pose estimation: A survey, 2021.

[221] Aojun Zhou, Anbang Yao, Yiwen Guo, Lin Xu, and Yurong Chen. Incremental network quantization: Towards lossless cnns with low-precision weights, 2017.

[222] Xiangyu Zhu, Yingying Jiang, and Zhenbo Luo. Multi-person pose estimation for posetrack with enhanced part affinity fields. In *ICCV PoseTrack Workshop*, volume 7, 2017.