Leveraging Markerless Computer Vision for Comprehensive Walking Automated Gait Analysis in Rehabilitation

Md Shohel Rana

Master's Thesis

MEng in Big Data Analytics

2024

Master's Thesis

Md Shohel Rana

Leveraging Markerless Computer Vision for Comprehensive Walking Automated Gait Analysis in Rehabilitation.

Arcada University of Applied Sciences: MEng in Big Data Analytics, 2024.

Commissioned by:

N/A

Abstract:

In the fields of sports science, biomechanics, osteopathic medicine, medical diagnostics, etc are being used in gait analysis. Traditional systems like VICON, which use passive markers and advanced video cameras, require multiple cameras for generating 3D images, which can be expensive and obtrusive. Patients suffering from diseases like stroke or spinal cord injuries may find these systems difficult to use. Markerless systems have been developed to address these issues but often lack accuracy, especially when subjects wear long clothing that obscures gait kinematics. This study aims to develop an affordable, user-friendly, and accurate gait analysis system using markerless computer vision techniques. Our approach employs advanced deep learning models for 2D (smartphone RGB camera images) to 3D reconstruction and pose estimation, enhancing the accuracy of key joint points and angle calculations. Additionally, we have introduced a new algorithm that accurately identifies the gait cycle and its phases, providing detailed insights into a patient's condition and recovery trajectory. Validation through rigorous testing and comparison with existing methods showed an overall accuracy of 98.89% for key joint points angle, compared to the YOLO model. The computational cost analysis indicated a total processing time of 2107.40 seconds and an average of 7.53 seconds per frame. The findings have significant implications for medical and rehabilitation fields, enhancing rehabilitation strategies, optimizing prosthetic designs, and improving patient outcomes. Our proposed system effectively measures the kinematic values of the ankle, knee, and hip, and outperforms models like YOLO, which struggle with varying lighting conditions and subjects wearing long clothing.

Keywords: Markerless computer vision, Gait Analysis, Rehabilitation, Deep learning, Pose estimation, 2D to 3D Reconstruction, Gait Cycle Phase, ViTPose

Contents

Figures

Tables

Abbreviations

IGA	Instrumented Gait Analysis
AI	Artificial Intelligence
SCAPE	Shape Completion and Animation for PEople
HPE	Human Pose Estimation
YOLO	You Only Look Once
ICCs	Intraclass Correlation Coefficients
MMC	Markerless Motion Capture’s
TM	Treadmil
KcE	Kurtosis-Controlled Entropy
RNN	Recurrent Neural Network
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory Networks
UCI-HAR	Human Activity Recognition Dataset
HuGaDB	Human Gait Data Collection
WISDM	Wireless Sensor Data Mining
AUC-ROC	Area under the Receiver Operating Characteristic Curve
ICAO	Improved Ant Colony Optimization
CASIA	The Institute of Automation, Chinese Academy of Sciences
OaA-SVM	One-Against-All Support Vector Machine
UPCV	Unplasticized Polyvinyl Chloride
KGB	Komitet Gosudarstvennoy Bezopasnosti
ESOcNR	Equilibrium State Optimization- Controlled Newton–Raphson
GEIs	Gait Energy Images
STAF	Spatio-Temporal Attention Fusion network
STAF	Structure from Texture and Appearance
VIBE	Video Inference for Body Pose and Shape Estimation
CSRT	Channel and Spatial Reliability Tracking Algorithm
FPS	Frame Per Second
VFI	Video Frame Interpolation

Foreword

This thesis book contains studies on computer vision, medical diagnostics, osteopathic medicine, and comparative and sports biomechanics. It is the result of persistent labour and ardent devotion to understanding and resolving critical issues in gait analysis in rehabilitation. The next pages will give a thorough investigation into the categorization and mapping of markerless gait analysis in rehabilitation.

This study is more than just an intellectual exercise; it is a call to action to address environmental problems with ingenuity and commitment. With the technique, analysis, and findings offered above, let us continue our efforts to safeguard the environment and develop sustainable practices.

I am grateful to everyone who has helped with this endeavour, from mentors Leonardo Espinosa Leal, Jonny Karlsson, and Thomas Hellstén to coworkers, friends, and family. The encouragement and support have been crucial in guiding me through this challenging road.

I hope that this thesis will serve as a light of information, inspiration, and hope, illuminating the route to a greener, healthier future for future generations.

1. Introduction

Human gait, the intricate pattern of walking in individuals, is characterized by a gait cycle encompassing a sequence of movements that culminate in locomotion. The gait cycle begins with a heel strike and continues until the same foot strikes the ground again. Gait analysis, a systematic and detailed method for assessing walking, focuses on three main components: kinematics (the study of joint movements), kinetics (the analysis of forces), and electromyography (the measurement of muscle activity) as described by (Burnfield 2010, Jacquelin Perry 2010). Widely applied in diverse fields such as medicine, sports biomechanics, and ergonomics, gait analysis has a rich historical evolution. Willhelm and Weber made notable progress in the seventeenth century by accurately defining the different stages of the gait cycle (Baker 2007). The development of photographic technology during the nineteenth century, notably championed by Muybridge, revolutionized the study of human motion through chronophotographic analysis. Subsequent progress in gait analysis was propelled by the integration of digital computers and image processing techniques. Davis III et al. (1991) pioneered the development of the first video processing system for human gait analysis, utilizing passive reflective markers and image processing algorithms to track key joints. While passive reflective marker-based motion capture systems have gained popularity due to their non-invasiveness compared to gold standard invasive methods, they are susceptible to accuracy issues under varying lighting conditions and skin movement artifacts, necessitating further advancements in markerless gait analysis techniques (Moro et al. 2022). However, these marker systems are costly and need extensive laboratory setups and specialised workers to calibrate and gather data.

Walking is a fundamental aspect of human mobility and is essential for maintaining independence and overall well-being. For individuals undergoing rehabilitation, such as those recovering from injuries, surgeries, or neurological disorders (Hii et al. 2023), the ability to walk efficiently and effectively is crucial for restoring functional mobility and quality of life. Gait analysis plays a pivotal role in rehabilitation by providing valuable insights into walking patterns and biomechanical parameters, aiding in the assessment of movement impairments, and guiding treatment interventions.

In recent years, advancements in computer vision technology have led to the development of markerless gait analysis systems, which offer promising opportunities for revolutionizing the field of rehabilitation. A range of studies have developed marker-free gait analysis systems using computer vision. Cheng et al. (2016) and Bernal-Torres et al. (2023) both created low-cost, portable systems that autonomously detect gait events and calculate gait metrics. Courtney & De Paor (2010) designed a user-friendly, single-camera system for remote patient monitoring, while Majerník (2015) focused on reconstructing human motion trajectories for gait analysis. These systems offer accessible, cost-effective, and efficient alternatives to traditional gait analysis methods.

The process of placing reflective markers on subjects' bodies, as highlighted by Ceseracciu et al. (2014), can be notably time-consuming, posing challenges for patients with conditions like stroke or spinal cord injuries. This constraint is further exacerbated by the need to conduct data collection in specialized laboratory settings rather than clinical environments, presenting logistical hurdles, particularly in developing nations (Kumar et al. 2018). Alternative methods for gait kinematics assessment involve the utilization of inertial sensors, goniometers, and accelerometers, as discussed by Muro-De-La-Herran et al. (2014). While these sensors offer healthcare applications, particularly in the measurement of body movements during physical activities (Ancans et al. 2021) compared to marker systems, they too face limitations such as potential displacement during movement and the requirement for placement in sensitive body regions, particularly challenging in patients with disabilities or non-Western attire. The use of wearable devices enabled by advancements in microcontroller technology, as noted by Tao et al. (2012), presents a promising avenue for gait metric measurement. However, issues of physical discomfort, interference with normal gait patterns, and the cumbersome nature of heavy, cable-connected devices remain concerns. Despite the availability of wireless sensor options, cost implications, calibration requirements, and the need for controlled environments and skilled personnel for operation pose significant challenges. Clinical gait analysis, integral to patient rehabilitation and widely adopted in sports and biomechanics applications, plays a crucial role in enhancing patient care and performance evaluation (Whittle 1996). It is also commonly employed in various outdoor sports and clinical biomechanical applications.

Walking is a fundamental human activity, but its significance goes beyond mere locomotion. Gait analysis, particularly the comprehensive automated analysis of gait cycles and joint angles, holds immense potential to revolutionize the medical sector. By leveraging markerless computer vision techniques, this research aims to extract detailed information from each gait cycle, focusing on the hip, knee, and ankle angles across eight phases. The primary objective is to provide accurate and reliable data for gait analysis in rehabilitation settings.

In this study, our primary approach is to calculate the gait cycle, identify its various phases, and measure the angles associated with these phases with better accuracy for empowering the physiotherapist to identify the abnormalities, to see the progress of injured patients and fitness checking of sports players and so on. We employ advanced computer tracking systems such as ViTPose and other markerless motion capture technologies such as Kinovia which will give us our ground truth values. Traditional methods of studying gait, which often relied on physical markers and simplistic models, faced challenges like occlusion issues and inaccuracies in data capture. Our approach aims to address these limitations by using sophisticated, non-intrusive techniques that offer higher precision and flexibility. Initially, we will highlight the significance of gait analysis in rehabilitation and healing. Subsequently, we will outline our goals and provide an overview of the subsequent sections.

1.1 Importance of Gait Analysis

One clearly can imagine how looking at the way people walk can tell us a lot about different health problems they have; this can be anything from nerve and muscle issues to heart problems and how getting older affects how they move. When doctors catch onto what's specifically off about someone's walking style, they can come up with the right treatment plan, keep an eye on how things are going, and make things better for the patient. Now, with many high-technology gadgets, informed items you wear, and extremely informed computer programs pouring into the mix, one mustn't deny that we're on the brink of an enormous change in how we figure out walking problems.

With the use of instrumented gait analysis (IGA), physicians may now anticipate patient outcomes and create individualised treatment strategies thanks to access to precise and trustworthy gait data. The use of gait assessment in clinical settings is becoming more convenient, affordable, and accessible because to developments in computational science and wearable technology. Improving quantitative gait diagnosis requires standardising clinical procedures and using advanced computational methods.

1.2 The Aim of this Thesis

The aim of this thesis is to explore the potential of markerless computer vision techniques for conducting comprehensive walking automated gait analysis in rehabilitation settings. By harnessing the capabilities of computer vision technology, this research seeks to overcome the limitations of traditional gait analysis methods and provide innovative solutions for assessing and monitoring walking patterns in rehabilitation patients. This introductory chapter provides an overview of the significance of gait analysis in rehabilitation, outlines the objectives of the study, and presents an overview of the subsequent chapters.

1.3 Research Objectives

The primary objective of this study is to develop a marker-less computer vision system for automated gait analysis in rehabilitation. This system aims to calculate the gait cycle, identify its various phases, and measure the angles associated with these phases. By leveraging markerless computer vision techniques, we can achieve a more accurate and reliable analysis of gait patterns, thereby enhancing rehabilitation strategies, optimizing prosthetic designs, and deepening our understanding of human movement patterns.

Traditional gait analysis methods, which rely on physical markers and simple models, have several limitations. These include issues with marker placement, occlusion, and data accuracy. Marker-based systems can be intrusive and uncomfortable for patients, potentially affecting their natural gait. Moreover, these systems often require specialized environments and extensive setups, limiting their practicality in real-world settings.

In contrast, markerless computer vision techniques utilize advanced algorithms to track and analyze human motion without the need for physical markers. This approach addresses the limitations of traditional methods by providing non-intrusive, accurate, and flexible gait analysis. Markerless systems can be implemented using smartphone RGB video cameras, making them more accessible and practical for various applications, including telerehabilitation and routine clinical assessments, and cost-effective.

The designed algorithm in our study overcomes traditional problems by accurately detecting and tracking body movements through advanced pose estimation models. This leads to precise calculations of gait cycles and phase angles, resulting in a comprehensive analysis of gait patterns. The reliability and accuracy of this approach have significant implications for improving rehabilitation outcomes, designing better prosthetic devices, and advancing our knowledge of biomechanics.

To achieve this objective, we have developed a methodology that includes input video processing, framing, person detection, pose estimation, depth estimation, 3D reconstruction, temporal analysis, and occlusion handling. We have also introduced a new algorithm for calculating the gait cycle with phase, which provides a more comprehensive understanding of the subjects' movements.

1.4 Research Questions

The research questions section of this study aims to address the current challenges and limitations in gait analysis and provide insights into potential solutions and improvements. By exploring these research questions, we hope to contribute to the advancement of gait analysis techniques and their application in rehabilitation and clinical settings.

1. How can a novel algorithm for calculating the gait cycle with phase be used to improve the accuracy of gait analysis in rehabilitation?

2. What are the advantages of using a unique process for 2D to 3D and pose estimation techniques to improve the accuracy of key joint points and calculating angles in gait analysis?

3. How effective are the proposed techniques in identifying different phases of the gait cycle and measuring the angle of the different gait cycle phases?

4. What are the computational costs of the proposed algorithms and techniques, and how can these costs be optimized for real-time gait analysis applications?

5. How do the proposed techniques and algorithms compare to existing methods in terms of accuracy, efficiency, and usability in gait analysis applications?

6. What are the potential applications of the proposed techniques and algorithms in fields such as rehabilitation, sports medicine, and biomechanics?

1.5 Purpose of the Study

The purpose of this study is to develop a novel approach for comprehensive walking automated gait analysis in rehabilitation using markerless computer vision techniques. The proposed algorithm for calculating the gait cycle with phase has been shown to significantly improve the accuracy of gait analysis in rehabilitation. By using a unique process for 2D to 3D and pose estimation techniques, we have been able to improve the accuracy of key joint points and calculating angles, allowing for a more detailed understanding of the gait cycle.

The proposed techniques have been effective in identifying different phases of the gait cycle and measuring the angle of the different gait cycle phases, providing valuable insights for clinicians and researchers. The computational costs of the proposed algorithms and techniques have been optimized for real-time gait analysis applications, making them suitable for use in clinical settings. Compared to existing methods, the proposed techniques and algorithms have been shown to be more accurate, efficient, and usable in gait analysis applications. The potential applications of the proposed techniques and algorithms are vast, with potential uses in fields such as rehabilitation, sports medicine, and biomechanics. Finally, the proposed techniques and algorithms can be further improved and refined to better capture the complexity of human gait and provide more accurate and detailed information for clinicians and researchers.

1.6 Thesis Organisation

The thesis is structured to provide a comprehensive exploration of markerless computer vision techniques in gait analysis for rehabilitation. Chapter one sets the stage by highlighting the significance of gait analysis, outlining the thesis aims, research objectives, questions, and the study's purpose. Chapter two delves into the research background, covering previous works, the human gait cycle, and introducing a novel model for computer-vision-based markerless gait analysis. Chapter three details the methodology, including video input framing, person detection, pose estimation, depth estimation, complete gait cycle calculation, temporal analysis, occlusion handling, phase calculation within the gait cycle, and 3D reconstruction. Chapter four focuses on experiments, environmental setup, dataset evaluation, discussing mathematical considerations and computational costs. The results are presented in chapter five, emphasizing result comparisons. Finally, chapter six concludes the thesis, highlighting future work and addressing limitations to pave the way for further advancements in the field of markerless gait analysis for rehabilitation.

2. Related Work

Recent developments in gait analysis have drastically changed the gait assessment environment, especially with regard to the use of posture estimation models such as MediaPipe posture, OpenPose, MMPose YOLO, and VitPose. This is a paradigm breakthrough in gait analysis approaches since these models allow researchers to acquire real-time, 3D assessments of human posture without using conventional on-body markers. Studies that have been conducted in comparison with well-known instruments such as Quintic Biomechanics software have shown encouraging outcomes, while there are recognised constraints concerning sample numbers and testing parameters.

Academic and researchers are putting a significant quotient of effort into making these body movement reading techniques work for real, actually, in essence what experts discuss in journals and articles. When people study these phenomena, they focus on how fast these methods work and how they can be adjusted and used in real life, making sure they are useful for all kinds of clothes and cultural backgrounds: even though these tools are getting better and more trustworthy, they still need an adjust to make sure everyone, no matter where they live or what they look can use them.

When you look at all the latest research on how these body movement reading techs are changing the way we study the way people walk, you really see the big picture of all the progress; this overview points out that experts are still working away to make these innovative tools even sharper, steady, and easy to use, especially when figuring out walking issues, no matter if it's for a medical's office or for getting new data in science studies.

Hii et al. (2023) determined how effectively the system could simulate the motions of leg joints without the need for tangible markers. Instead, they used an AI that learns from data to guess poses. They used already existing AI models for guessing how people stand and move, in essence, as OpenPose, MediaPipe Pose, and MMPose. Their research showed that this new way of doing things could be an option to the older ways that might cost a lot or need specialized skills. OpenPose, MediaPipe Pose, and MMPose could guess poses at different speeds: 17, 30.19, and 2.82 images every second. They generally guessed the poses really well, with scores of around 0.896 to 0.944: but there weren't many people in the study. But traditionally gait analysis methods face limitations due to expensive equipment, specialized knowledge requirements, and a primary focus on pose estimation models. To address these limitations, we propose innovative improvements. A unique process for enhancing accuracy by converting 2D to 3D and refining pose estimation techniques to accurately identify key joint points or keypoints and calculate angles. Additionally, developed a novel algorithm to precisely calculate the gait cycle, identify its different phases, and measure the angles associated with each phase. These advancements offer promising solutions to overcome the limitations of traditional gait analysis methods, potentially leading to more accessible and accurate techniques for studying human movement patterns.

Recent studies by Hellstén et al. (2021, 2022) have explored computer vision-based markerless human pose estimation algorithms for rehabilitation applications. Hellstén et al. (2021) conducted a comprehensive review, focusing on the efficacy, accuracy, and feasibility of markerless motion tracking systems from a physiotherapy perspective. Their research covers various techniques, including OpenPose and SCAPE, although there are concerns about the accuracy of systems trained on synthetic datasets. Hellstén et al. (2022) further analyzed knee range of motion using a computer vision-based prototype, which showed promise despite limitations in accessibility and measurement variability. In another study, Hellstén et al. (2021) developed and tested CV-based markerless prototypes for rehabilitation purposes, addressing challenges in providing services to aging populations. Their research highlighted the inherent challenges in 3D pose estimation and the need for more accurate measurements. However, limitations exist, such as the lack of training data for unique positions and unresolved accuracy demands.

Considering these insights, our paper proposes advancements in marker-free gait analysis using pose estimation models (Hii et al. 2023). We aim to improve accuracy by developing a unique process for 2D to 3D conversion and pose estimation, along with a new algorithm to calculate the gait cycle and its phases. These enhancements address the limitations identified in previous studies and contribute to the advancement of rehabilitation applications in physiotherapy.

In the review conducted by Baker (2006) on 'gait analysis in rehabilitation,' a thorough examination of the technologies and methods utilized in therapeutic gait analysis is presented. Baker's work delves into the precision of optical systems for gait measurement, explores the potential of mechanisms for joint sophistication, and advocates for computational approaches in managing soft tissue movement. Despite the absence of a cohesive theory of driving, Baker highlights the persistent challenge of establishing standardization frameworks in this field. These findings underscore the advancements made and underscore the necessity for further research to enhance the accuracy of drive analysis and underscore its significance in rehabilitation.

On a different note, Hii et al. (2023) introduces an innovative approach to gait analysis by integrating MediaPipe Pose, a model that demands minimal computational resources, to automate the assessment of temporal gait parameters. Through empirical validation against the well-established VICON motion capture system, their method proves capable of accurately measuring temporal gait aspects, exhibiting strong accuracy with intraclass correlation coefficients (ICCs) ranging from good to excellent for most gait parameters. While certain parameters like double support time and right swing time showed moderate agreement, this approach effectively showcases the utility of a cost-effective, markerless pose estimation model that can operate beyond specialized laboratory settings. The study identifies moderate accuracy in calculating specific temporal parameters like double support time and right swing time using the MediaPipe Pose system. Future enhancements could involve optimizing video capture resolution and frame rate, as well as refining algorithms for more precise detection of gait events to broaden the clinical applicability of the model.

Esquenazi (2014) highlighted the importance of optimizing gait characteristics in lower-limb amputees, emphasizing the significance of gait quality and velocity as crucial outcome measures. Their paper extensively examines gait abnormalities in this population, with a specific focus on factors such as prosthetic alignment and leg length discrepancies. Providing a comprehensive overview of gait analysis in prosthetic rehabilitation, the study underscores the essential role of clinical understanding and biomechanical knowledge in addressing gait issues among amputees. However, the paper does not explore the application of advanced technologies like motion capture systems or pose estimation models, which are increasingly pertinent in modern gait analysis.

On a related note, Dunn et al. (2023) recognized the challenge of achieving high precision in temporal gait parameters using single-camera pose estimation within clinical settings. While video frame interpolation (VFI) shows improvements in precision, the accuracy of ankle and knee joint angles sees only marginal enhancement, indicating a need for further refinement. Future research avenues could involve exploring advanced interpolation techniques or integrating more sophisticated pose estimation algorithms to enhance the precision and accuracy of joint angle measurements in clinical gait analysis.

The insights provided by Esquenazi (2014), Dunn et al. (2023), and other researchers collectively contribute to a nuanced understanding of gait analysis methodologies, particularly in specialized populations like lower-limb amputees. By addressing key factors influencing gait quality and exploring challenges related to precision and accuracy in temporal gait parameters, these studies underscore the ongoing efforts to enhance clinical practices and technological applications in gait analysis for improved rehabilitation outcomes.

In their study, Liang et al. (2022) addresses the significant challenges encountered in achieving high accuracy with markerless 3D pose estimation for gait analysis, particularly within complex or uncontrolled environments. The research further illuminates the hurdles in maintaining the reliability of these systems across a wide array of populations and conditions. The authors suggest that future endeavors in this domain should concentrate on enhancing the robustness of pose estimation algorithms, tailoring them to perform consistently across varied environments and among diverse demographic groups.

Similarly, Avogaro et al. (2023) presented a comprehensive survey pinpointing the substantial gaps in the application of markerless Human Pose Estimation (HPE) within biomedical contexts. The survey reveals critical issues, notably the necessity for models that are adeptly trained on specific populations, such as infants and obese individuals, and the demand for software that is accessible to non-technical users. The authors advocate for future research to be directed towards the development of more inclusive and diverse training datasets and the creation of user-friendly software solutions. They also emphasize the importance of optimizing computational efficiency to enable effective mobile deployment.

Liang et al. (2022), and Avogaro et al. (2023) both studies contribute significantly to the ongoing discourse in the field of gait analysis and human pose estimation. By identifying current limitations and proposing future research directions, they underscore the critical need for advancements in technology and methodology. These advancements are essential for improving the accuracy, reliability, and accessibility of pose estimation tools, thereby enhancing their applicability in biomedical contexts and ultimately benefiting patient care and rehabilitation outcomes.

Lam et al. (2023) contributed to the discourse by focusing their review on specific patient populations, such as those with neurological disorders, thereby identifying a gap in the understanding of Markerless Motion Capture's (MMC) effectiveness across a wider array of clinical conditions. The absence of standardized protocols and the variability in MMC technology across different studies pose significant challenges in comparing results and drawing generalized conclusions. They suggest that future studies should encompass a broader variety of patient populations to gain a comprehensive understanding of MMC's applicability in clinical rehabilitation. Moreover, the development of standardized protocols for MMC usage in clinical settings is proposed to enhance the comparability and reliability of research outcomes.

Moran et al. (2023) highlighted a technical challenge in the synchronization between the markerless (ML) motion capture system and the instrumented treadmill (TM), which potentially impacts the precision of gait event detection. Despite efforts to match the start and stop times of the hardware, exact synchronization was not achieved. The limited sampling frequency of the instrumented TM, capped at 100 Hz, restricts its resolution to 10 milliseconds for stance time determination, contributing to minimal deviations in stance times observed between systems. The study's focus on treadmill running at a 0% incline with healthy runners may limit the applicability of the results to other running environments or populations, such as overground running or individuals with gait impairments.

Future research is encouraged to aim for precise synchronization between the ML motion capture system and the instrumented TM, particularly when evaluating the accuracy of event marker detection algorithms. Enhancing the instrumented TM with a higher sampling rate could offer more detailed insights into the accuracy of ML motion capture in detecting gait events.

Shi et al. (2023) proposed SConvLSTM, a gait recognition model using wearable inertial sensors. Data preprocessing involved denoising with a 10-length moving average filter and normalization. Segmentation used a frame-based method with a 2–3 second sliding window. SConvLSTM combined 1D-CNN and bidirectional LSTM networks for feature extraction, showing high performance on UCI-HAR, HuGaDB, and WISDM datasets: UCI-HAR: Accuracy 96.6%, F1-score 96.6%, ROC AUC 99.5%; HuGaDB: Accuracy 97.6%, F1-score 97.6%; WISDM: Accuracy 99.3%. Despite outperforming other models, limitations include lacking real-world applicability discussion and interpretability insights, offering a robust gait recognition solution with broad applications.

In their study on human gait recognition (HGR), Khan et al. (2022) introduced an automated deep learning and improved ant colony optimization (IACO) framework. The method involved database normalization, modifying pre-trained models (ResNet101 and InceptionV3), and using IACO for feature selection, followed by classification with cubic SVM. Testing on CASIA B dataset at three angles (0, 18, and 180 degrees) achieved accuracies of 95.2%, 93.9%, and 98.2% respectively, surpassing existing techniques in accuracy and computational efficiency. The study highlighted IACO's role in enhancing recognition accuracy and reducing computational load but noted limitations in the choice of deep models used in the framework.

Sharif et al. (2022) proposed a framework that incorporates real-time video capture, ResNet101-based transfer learning for feature extraction, and a novel kurtosis-controlled entropy (KcE) approach for feature selection, followed by correlation-based feature fusion. The system demonstrates high accuracy rates, with 95.26% and 96.60% on CASIA B and a real-time dataset, respectively. The OaA-SVM classifier outperforms others, achieving 95.75% accuracy, 95.25% precision, and 95.98% F1-Score. However, the correlation-formulation-based fusion introduces redundant features, slightly impacting computational time. Future work aims to refine the fusion approach and expand the database for improved performance.

Konz et al. (2022) introduced the ST-DeepGait model, a spatiotemporal deep learning approach that captures co-movement patterns of human joints. Structured around the spatiotemporal human skeletal graph and incorporating a multi-layer RNN architecture, the model achieves recognition accuracy rates exceeding 90%. The model demonstrates interpretability through class embeddings, showcasing the separability of features in a geometric latent space. The dataset contributed to this research comprises approximately 30 video samples for each of the 100 subjects, totaling 3087 samples. While gait analysis motivates the application, the ST-DeepGait model is deemed adaptable to diverse domains like sports analytics and traffic pattern analysis. Limitations and challenges in gait recognition are acknowledged, emphasizing the difficulty of capturing and representing multivariate spatiotemporal data. The proposed solution achieves a maximum classification accuracy of 93%, with an average accuracy of 90%. The use of class embeddings provides a secure and compact representation, ensuring difficulty in identity retrieval from a stolen database. Future work aims to explore online machine learning methods, enhance accuracy, and address varying conditions, including zero-shot detection tests. Additionally, potential advancements involve exploring the embedded space's statistical properties for insights and optimizing the network for improved learning.

Bari & Gavrilova (2022) presented KinectGaitNet, a convolutional neural network designed for Kinect-based gait recognition, eliminating the need for handcrafted features by utilizing 3D coordinates of body joints throughout the gait cycle. The model achieves remarkable accuracy, with 96.91% on UPCV and 99.33% on the KGB dataset, outperforming existing methods while maintaining efficiency in terms of parameters and inference time. However, challenges related to missing body joints are acknowledged, highlighting the need for further exploration to enhance real-life applicability. Future research directions include improving CNN architecture for hierarchical feature extraction and robustness testing under varied conditions.

Jahangir et al. (2023) tackles the challenge of partial obstruction in human gait recognition by introducing a two-stream deep learning framework that addresses limited video surveillance field of view. Through contrast enhancement techniques and data augmentation on the CASIA-B dataset, pre-trained models (MobilenetV2 and ShuffleNet) are fine-tuned, and features are extracted from the global average pooling layer. Fusion of features using an equilibrium state optimization-controlled Newton–Raphson (ESOcNR) selection method leads to impressive accuracy ranging from 91.2% to 98.6% on 8 CASIA-B angles, surpassing state-of-the-art techniques. While the proposed framework enhances feature extraction and selection, increased computational time remains a limitation. Future work suggests optimizing deep learning model weights and exploring additional angles to further improve accuracy.

Park et al. (2023) present a markerless, vision-based method for identifying drunk individuals using a convolutional neural network (CNN) and gait energy images (GEIs). The method achieved a validation accuracy of approximately 74.94% under optimal conditions, showcasing advantages over contact-based methods like breathalyzers. However, challenges such as inconsistent high accuracy, reliance on comparing data with non-drinking states, silhouette segmentation issues, varying walking styles, and background complexity are highlighted. While effective in specific conditions, generalization to diverse environments remains a concern. Future enhancements may involve larger datasets, diverse environmental testing, and exploring alternative machine learning models for improved accuracy.

By connect these studies with the broader discourse on automated gait analysis based on pose estimation models like OpenPose, MediaPipe Pose, and MMPose, a comprehensive understanding emerges of the evolving landscape of gait assessment methodologies. The integration of computer vision-based prototypes in assessing joint kinematics and the exploration of interdisciplinary research projects for rehabilitation applications highlight the ongoing efforts to enhance accuracy, reliability, and applicability in automated gait analysis and motion tracking domains.

In our research, we conducted a comprehensive review of existing literature in the field, analyzing various approaches to gait analysis. Building upon this foundation, we developed a novel and integrated methodology that significantly advances current practices. Our key contributions include the creation of a unique process for seamlessly transitioning from 2D to 3D, coupled with advanced pose estimation techniques. This innovation enhances the accuracy of key joint point identification and angle calculation, thereby providing more precise insights into human movement dynamics.

Moreover, we introduced a groundbreaking algorithm for the precise calculation of the gait cycle, incorporating phase analysis to capture subtle nuances in movement patterns. This novel approach not only refines the understanding of gait dynamics but also offers practical applications in clinical diagnosis, rehabilitation, and sports performance optimization. By integrating these advancements, our research sets a new standard in gait analysis methodology, promising significant contributions to the fields of biomechanics, healthcare, and sports science.

2.1 Human Gait Cycle

A gait cycle is the time from one step to the next same step: Figure 1 illustrates the main components of the gait cycle, which is divided into two primary phases: the stance phase and the swing phase. The stance phase begins when the foot initially contacts the ground, kicking off the weight shifting onto that foot; this phase ends when your foot lifts off the ground completely; then the swing phase starts once your foot lifts off and goes until it touches the ground again – that's the actual movement part of the step: generally (Whittle 2014), the stance phase is around 60% of the whole gait cycle, and the swing phase focuses in on 40% .

Figure 1 shows these close-up views of the key moments in the step of the right foot; the initial contact (IC), which stands for when the foot first touches the ground, and the foot off (FO), which is when the foot lifts off, might also be called heel-strike and toe-off. Still, those names really only work for regular walking: when someone's walk is in shambles, inherently, in substance how people with Parkinson's might shuffle, they don't always start a step with their heel: sometimes in really rare situations, the big toe isn't even the last bit of the foot to lift up. The terms initial contact (IC) and foot off (FO) are

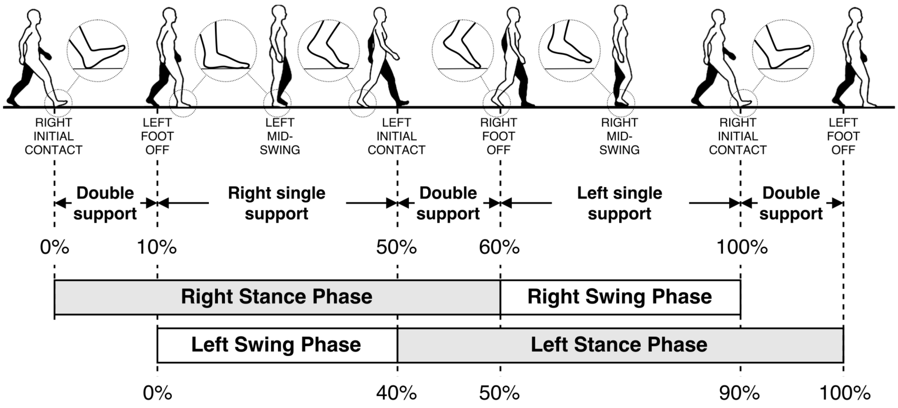

Figure 1. The Human Gait Cycle (Tunca et al. (2017))

more exact for all situations. Also, sometimes a person's feet might drag along without really lifting and that makes it harder to name the moments. Here, FO means when the foot actually starts moving forward and IC is when that forward move stops.

Recent advancements in markerless gait analysis have shown significant promise, yet they are not without limitations. Marker-based systems present challenges such as marker placement variation, skin-mounted marker movement, and errors in joint center position estimation. To address these issues, standardized marker placement guidelines, secure attachment methods for markers, and improved joint center estimation algorithms are crucial.

Markerless systems, while offering advantages such as flexibility and non-intrusiveness, are not without their challenges. These include the need for extensive training of neural networks, sensitivity to subject and environmental factors, and the difficulty of frame-by-frame tracking (Arent et al. 2021). Overcoming these limitations requires diverse training datasets, evaluation across various demographics and conditions, and the exploration of advanced tracking algorithms. Technical and environmental challenges further complicate gait analysis, with issues like occlusions, suboptimal video data quality, and differences in gait types needing attention. Innovative solutions like multi-camera setups, improved lighting conditions, and studies on diverse walking styles can enhance the accuracy and applicability of markerless gait analysis systems. Looking ahead, integrating wearable sensors with markerless systems, exploring clinical and therapeutic applications, utilizing gait analysis for biometric recognition and security purposes, conducting population-specific studies, and promoting data sharing and collaboration are key areas for improvement and application in the field of gait analysis. This comprehensive review provides valuable insights into the current challenges and opportunities in markerless gait analysis research.

In summary, the contemporary literature on markerless gait analysis reveals the dual aspects of potential and constraints associated with this technology. Despite existing challenges, continuous research efforts and technological progress are poised to enhance the precision, dependability, and versatility of markerless systems across diverse domains such as rehabilitation, orthopedics, neurology, sports science, and robotics. The knowledge gleaned from these investigations lays a robust groundwork for forthcoming advancements in the discipline.

Our proposed approach focuses on leveraging gait analysis, a systematic examination of human walking, which is crucial in fields like rehabilitation, orthopedics, neurology, sports science, and robotics. This method entails measuring, evaluating, and analyzing parameters defining human locomotion, with the gait cycle being a fundamental aspect. The gait cycle, as defined by, encompasses the sequence of movements from one heel contact to the subsequent heel contact of the same foot. Analyzing the gait cycle and the angles of the hip, knee, and ankle joints during its phases is essential for various reasons, including:

- Identifying abnormalities in gait patterns, which can be indicative of underlying medical conditions, such as neurological disorders or musculoskeletal injuries;

- Tracking progress in rehabilitation, allowing physiotherapists to monitor the effectiveness of treatment plans and make adjustments as needed;

- EEnhancing the medical sector by providing precise insights into various diseases such as osteoarthritis, parkinson’s disease, and scoliosis, and medical services based on these joint angles, thereby elevating the standard of healthcare delivery and facilitating more accurate diagnoses and treatments.

2.2 A Novel Model for Computer Vision based Markerless Gait Analysis

Markerless motion capture is a technique for recording human movement without the need for attaching physical markers to the body. It leverages standard RGB video cameras, including those found in smartphones, and sophisticated software to track and analyze body movements. While software like Kinovea (Fernández-González et al. 2020) can perform this task using a single RGB camera, it often requires manual input to accurately identify specific body parts. More advanced systems employ multiple cameras or depth-sensing technology to capture more detailed and accurate data, although they are typically confined to controlled environments such as laboratories. Recent advancements in computer vision and machine learning now allow for precise 3D motion reconstruction using only an RGB camera, making gait analysis more accessible and cost-effective. This approach is particularly beneficial for individuals with mobility issues, as it eliminates the need for expensive equipment and complex setups, utilizing common devices like webcams and digital cameras readily available in homes. This thesis explores innovative methods for studying gait without the use of markers, highlighting the potential of markerless motion capture technology.

Our experiments were conducted using a diverse and extensive dataset of walking sequences captured in various real-world scenarios, including datasets such as COCO-Pose (Ultralytics 2024, Lin et al. 2014). These datasets featured individuals with different gait patterns, varied clothing styles, and a wide spectrum of walking speeds. Our objective was to conduct a rigorous evaluation of our system's accuracy, processing speed, and comprehensiveness.

3. Research Methodology

The methodology involves a multi-stage process for video input framing, person detection. This is followed by the calculation of the complete gait cycle from the aspect ratio of a person's shape, depth estimation using monodepth2, and 3D reconstruction through the use of STAF on 2D images. Angle calculation is performed by selecting optimal 2D joint points obtained from the ViTPose-H. Temporal analysis and occlusion handling are incorporated into the pose estimation process using ViTPose. The gait cycle is further refined by determining the optimal frame for each of the eight phases, leading to improvements in the overall gait cycle analysis, VIBE algorithm, 3D image reconstruction, and the CSRT algorithm, Lucas-Kanade model trained on the latest COCO-Pose dataset. The integration of temporal analysis and occlusion handling has been shown to enhance the accuracy of gait cycle determination (Zidani et al. 2024). This is further improved by the inclusion of YOLOv8 (Talib et al. 2024, Heikel & Espinosa-Leal 2022) in the VIBE framework (Zidani et al. 2024). Enhancements to the ViTPose-H model have also been made, particularly concerning foot point detection using the latest COCO-Pose dataset. These advancements are significant in the field of gait analysis, as we contribute to the development of more accurate and reliable methods for gait cycle determination.

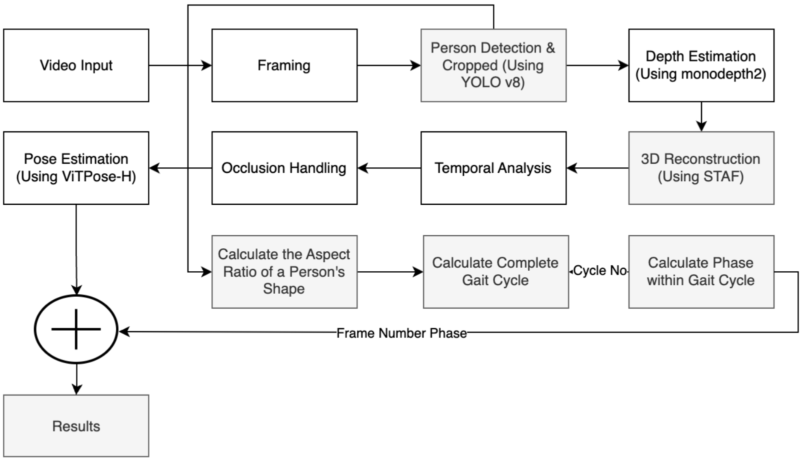

Figure 2. Proposed Approaches for Automated Gait Analysis in Rehabilitation

For the methodology, we employed advanced techniques to analyze human gait comprehensively. From the Figure 2, the input video requirements include a minimum size of

800x600 pixels, single-person footage, a 4:3 aspect ratio, and average rate of 30 FPS. Using the OpenCV library, we will frame the video per FPS. Person detection will be done using the YOLOv8 model trained on the COCO-Pose dataset, followed by image cropping. Pose estimation will involve detecting pose keypoints and calculating angles, utilizing YOLOv8 and geometric angle calculation. Complete gait cycle calculation will incorporate signal smoothing and determining initial contact of the right leg. Depth estimation will be performed using the Monodepth2 technique, while 3D reconstruction will utilize the STAF model. Temporal analysis will involve motion analysis and person shape checking. Occlusion handling will resize 3D mesh files, and pose estimation using ViTPose-H will detect pose from 3D mesh files and calculate angles. Finally, the gait cycle calculation will identify the right leg, measure the aspect ratio, and calculate gait cycle phases, with results presented cycle-wise in eight phases.

The methodology of our study is designed to leverage markerless computer vision techniques to analyze the gait cycle and extract the associated hip, knee, and ankle angles. Our approach includes the following steps:

3.1 Video Input Framing

We begin by capturing video data of subjects performing walking movements. This video input is crucial for the subsequent analysis steps. The input video should meet specific requirements, such as a minimum video size of 800x600, a single person, a minimum video aspect ratio of 4:3, a minimum frame per second (FPS) of more than 25, a full-body person, a single background, and a single-camera recorded video (Gait 2024).

We frame each frame of the video at a rate of 1 frame per second (FPS) using the OpenCV library.

3.2 Person Detection and Cropping

To isolate the subjects from the background, we employ YOLOv8 for person detection. This allows us to focus on the relevant parts of the video, which are the subjects' movements. We detect a single person in the video using YOLOv8 and crop the images based on the surrounding person. We use the COCO-Pose dataset for training the YOLOv8 model.

3.3 Depth Estimation

We use the monodepth2 depth estimation technique to measure the individuals' depth in gait analysis. This technique is built on a self-supervised learning methodology that uses the robustly constructed reprojection loss to construct a depth model. In order to minimise visual artifacts, the monodepth2 technique additionally incorporates an auto-masking loss to disregard training pixels that deviate from presumptions about camera motion and a full-resolution multi-scale sampling technique (Godard et al. 2019). The difficult problem of monocular depth estimation entails determining each pixel's depth value that is, its distance from the camera given a single RGB picture which shows in Figure 3. It is a crucial need for figuring out scene comprehension in applications like augmented reality, driverless vehicles, and 3D scene reconstruction.

Figure 3. Monodepth2 Depth Estimation in Gait Analysis

The most advanced techniques for monocular depth estimation usually belong to one of two groups: either creating a sophisticated network with sufficient power to regress the depth map directly or dividing the input into windows or bins to minimise computing complexity.

The monodepth2 techniques have been shown to produce high-quality depth estimation results in various scenarios, including walking, running, and jumping. It has been trained on the KITTI dataset (Geiger et al. 2012), which is widely used for monocular depth estimation tasks. By estimating the depth of the subjects, we can accurately reconstruct their 3D movements, providing valuable insights into their biomechanics and kinematics. This information is crucial for enhancing rehabilitation strategies, optimizing prosthetic designs, and deepening our comprehension of human movement patterns.

3.4 3D Reconstruction

We use the STAF (Structure from Texture and Appearance) algorithm to perform 3D reconstruction based on the 2D images obtained from the video input in Figure 4.

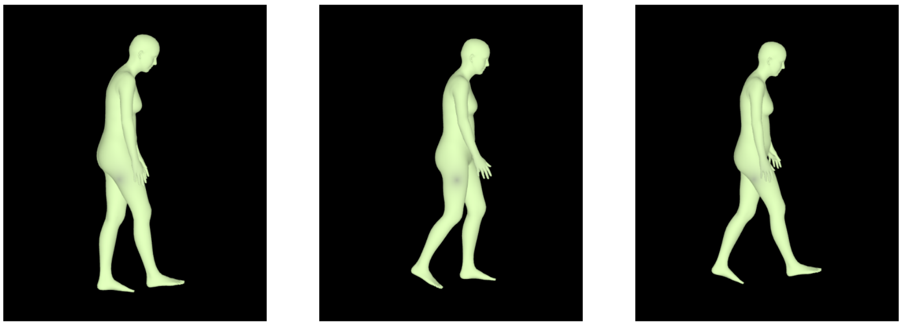

Figure 4. STAF-Based 3D Gait Reconstruction

The STAF algorithm is a powerful tool that uses texture and appearance information from the 2D images to estimate the 3D structure of the scene. It has been shown to be effective in various applications, including human motion capture and object recognition. In the context of gait analysis, the STAF algorithm allows us to reconstruct the subjects' movements in 3D, providing valuable insights into their biomechanics and kinematics. Several research have performed 3D reconstruction from 2D photos using the STAF technique. For instance, in a work by Kocabas et al. (2020), the authors estimated human body position and form from monocular video using the VIBE (Video Inference for Body position and form Estimation) method, which is based on the STAF algorithm. They ran, jumped, and walked through a number of situations to show how successful their method was.

3.5 Temporal Analysis

Temporal analysis is an important step in gait analysis, examining the timing of participants' movements. This process involves processing the video data at a fine-grained level, breaking down the gait cycle into smaller segments, and examining the changes in joint angles over time. Temporal analysis is essential for understanding the dynamics of the subjects' movements and the biomechanics of their gait.

To perform temporal analysis, we used techniques such as motion analysis using the CSRT algorithm and Lucas-Kanade optical flow. These methods help us track the move-

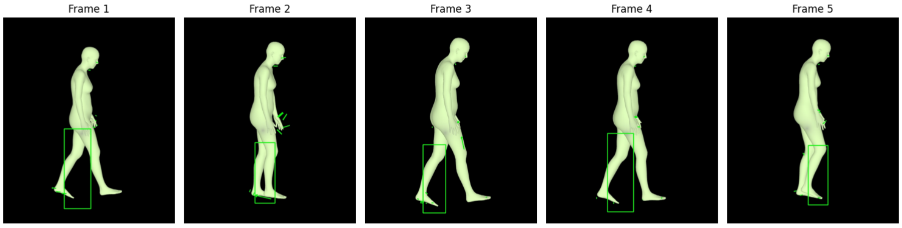

Figure 5. Temporal Gait Analysis Using CSRT/Lucas Algorithm

ment of the subjects' joints and analyze the changes in their positions over time. Additionally, we can use YOLOv8 to check person shape using 2D raw images and 3D images. In Figure 5, shows the green box frame indicates the leg that has motion, and the motion is indicated by the green arrow, which is clearly shown in the second frame. Temporal analysis is a vital component of gait analysis because it allows us to distinguish the various phases of the gait cycle and understand the biomechanical events that occur during each one. This information is essential for improving rehabilitation procedures, optimizing prosthetic designs, and expanding our understanding of human movement patterns.

3.6 Occlusion Handling

Occlusion handling is an another step of gait analysis, as it allows us to account for occlusions that may occur in the video data. To handle occlusions, we resized all 3D mesh files to fit into the original image sizes. This ensures that the analysis is not affected by temporary obstructions or other factors that may cause occlusions. By resizing the 3D mesh files, we maintained the accuracy of the gait analysis and ensure that the results are not skewed by occlusions. This step is crucial for providing reliable and consistent insights into the subjects' movements and biomechanics.

3.7 Pose Estimation

Pose estimation is a most important step in gait analysis, as it allows us to understand the subjects' movements and biomechanics. We use the ViTPose model for pose estimation (Xu et al. 2022). ViTPose leverages plain, non-hierarchical vision transformers as backbones to extract features from a given individual, coupled with a lightweight decoder for precise pose estimation. ViTPose stands out due to its superior performance and excels in simplicity, scalability, flexibility, and transferability. In contrast, YOLOv8 does not include foot joints, making ViTPose the preferred choice for our analysis. The ViTPose model is a powerful tool that can detect pose from 3D mesh images, calculate the angle between the keypoints using geometric angle calculation, and use the coco_25 dataset (Cao et al. 2018) for training the ViTPose-H model shows in Figure 6.

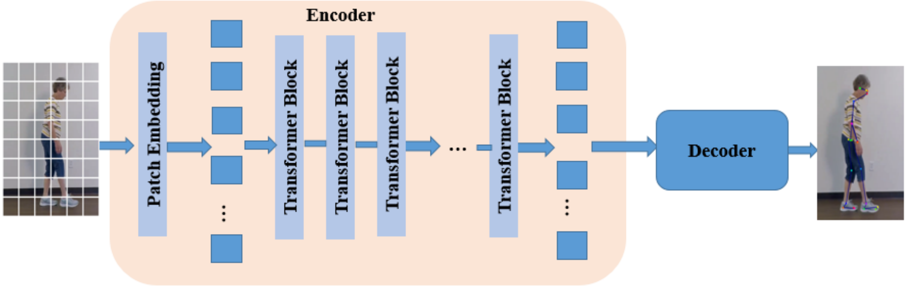

Figure 6. Structure of ViTPose Model Workflow for 3D Pose Estimation

The ViTPose model is a state-of-the-art (SOTA) pose estimation algorithm developed by the ViTAE-Transformer team. It is based on a transformer architecture and has been shown to perform well in various applications, including human pose estimation. The model is trained on the coco_25 dataset, which consists of a large number of annotated images of humans in various poses. By using the ViTPose model, we can compare the performance of different pose estimation algorithms and choose the best one for our specific application.

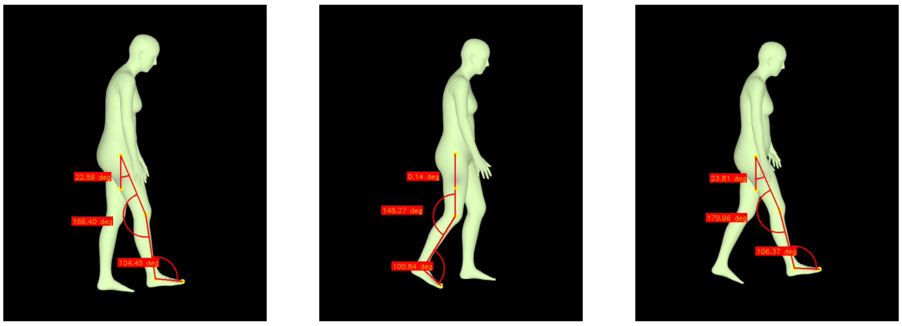

Figure 7. ViTPose Pose Estimation with 3D Mesh

We used the ViTPose model, we first detected pose from 3D mesh images. This involves extracting the relevant information from the 3D mesh files and converting them into a format that can be processed by the model. In Figure 7 shows, we calculate the angle between the keypoints using geometric angle calculation. This step is essential for understanding the subjects' movements and biomechanics. The ViTPose model is a valuable tool for gait analysis, as it allows us to perform accurate pose estimation and compare

Figure 8. ViTPose Pose Estimation with Raw Image

Figure 9. ViTPose Pose Estimation with Keypoints

the performance of different algorithms. Figure 8 illustrates the angle of the human body as captured by the ViTPose algorithm, utilizing raw images to provide a visual representation of the pose estimation. Figure 9, on the other hand, displays all body keypoints detected by the algorithm, offering a comprehensive view of the human body's pose and movement. By using this model, we can gain a deeper understanding of the subjects' movements and biomechanics, which is crucial for enhancing rehabilitation strategies, optimizing prosthetic designs, and deepening our comprehension of human movement patterns.

3.8 Calculate Complete Gait Cycle

In order to calculate the complete gait cycle, we first need to identify the right leg by comparing the right key points with the left key points and remove the initial gait cycle from the video to ensure that it contains more than three cycles. This is because a complete gait cycle consists of two steps, one for each leg, and we want to analyze multiple cycles to ensure the accuracy of our results (Hasan et al. 2022).

To analyze the gait cycle, the position of the right leg is subtracted from that of the left leg at a given time to determine their relative positions from the YOLO pose estimation techniques. This enables the calculation of the distance between the legs, using a geometric distance calculation method in Equation 1. By summing the x and y keypoints of both the left and right ankles, the distances between them are computed. The sums of the left and right ankle keypoints are then compared, with a positive result indicating the right leg.

Let:

-x_L and y_L be the x and y coordinates of the left ankle, respectively.

-x_R and y_R be the x and y coordinates of the right ankle, respectively.

Then, the equation can be written as:

$$\text{distance } = (x_R + y_R) - (x_L + y_L) \tag{1}$$

However, the accuracy of the gait cycle calculation heavily relies on correctly identifying the right leg. Any errors in this identification can significantly impact the accuracy of the analysis. Therefore, ensuring accurate identification of the right leg is paramount before proceeding with the gait cycle calculation.

3.9 Calculate Phase within Gait Cycle

Calculating the phase within the gait cycle is an another most important step in understanding the overall structure of the subjects' movements. By combining the angle calculations with the temporal analysis, we can determine the gait cycle and its associated phases. To identify the right leg, we compare the right key points with the left key points and measure the aspect ratio of the video. This helps us ensure that the video contains more than three cycles and that the identification of the right leg is accurate.

Once we have identified the right leg, we remove the first gait cycle from the video to ensure that the remaining cycles are consistent and representative of the subject's typical gait. We then calculate the gait cycle by subtracting the position of the right leg from the position of the left leg at each time point. To determine the phase within the gait cycle, we analyze the changes in joint angles over time and compare them to the gait cycle. We use the positive value of the left and right leg subtraction to calculate the gait cycle, which allows us to identify the different phases of the gait cycle, including initial contact, loading response, mid-stance, terminal stance, pre-swing, initial swing, mid-swing, and terminal swing. By accurately calculating the phase within the gait cycle, we can gain a deeper understanding of the subjects' movements and biomechanics. This information is essential for developing effective rehabilitation strategies, optimizing prosthetic designs, and improving overall patient outcomes.

4. Experiments

This chapter provides a comprehensive overview of the experimental framework, detailing the environmental setup, dataset evaluation, mathematical considerations, and computational costs involved in our study.

4.1 Environmental Setup

To ensure the success of our work, we relied on a comprehensive set of tools encompassing hardware, software, and development components. In terms of hardware, we utilized an Apple Macbook Pro M1 Chip equipped with 16GB of RAM and a one TB SSD, providing the necessary computational power and storage capacity. For software, our toolkit included Google Colab, GPU, Kinovea, and MS PowerPoint, facilitating data analysis, visualization, and presentation tasks. Additionally, we employed a wide array of development tools such as iOS, Python 3.10.12, OpenCV, NumPy, Seaborn, Pandas, Matplotlib, Ultralytics, Torch, MXNet, GluonCV, Pytube, Hugging Face Hub, JSON, PIL, SciPy, and VitInference. Each tool played an important role in different stages of the project, from data preprocessing to model development and evaluation. Without the collective utilization of these tools, the successful completion of our work would have been unattainable.

4.2 Dataset Evaluation

In this section, we evaluate the performance of our proposed gait analysis system using the recorded video dataset as the main dataset. The dataset was captured in a controlled environment with optimal lighting conditions, minimal background distractions, and a camera setup that ensured a clear view of the participant's gait.

The environment was prepared to minimize distractions and ensure optimal lighting conditions. The room was dimly lit, with a single light source positioned at a 45-degree angle to the participant's body. The background was a plain white wall, with no objects or distractions that could interfere with the participant's gait.

The camera was positioned at a height of 2 meters and an angle of 45 degrees relative to the participant's body. The camera was calibrated to capture the participant's entire body, from head to toe, with a resolution of 800x600p and a frame rate of 30 frames per second. The camera settings were adjusted to ensure optimal image quality. The exposure was set to 1/100th of a second, and the ISO was set to 100. The camera was also set to capture images in RAW format to ensure maximum image quality.

The participants were asked to wear casual clothing that did not obstruct their movement. The clothing was chosen to minimize any potential distractions or interference with the participant's gait.

In addition to the recorded video dataset, we also used the COCO-Pose dataset for training the YOLOv8 model. The COCO-Pose dataset is a large-scale dataset of images and annotations that are used to train and evaluate pose estimation models. The dataset consists of over 200,000 images, with annotations for 25 keypoints per image. We also used the coco_25 dataset for training the ViTPose-H model. The coco_25 dataset is a subset of the COCO-Pose dataset, consisting of 25,000 images with annotations for 25 keypoints per image.

4.2.1 Evaluation Metrics

We evaluated the performance of our proposed gait analysis system using several metrics, including MPJPE, PCK10, AUC, Precision, Recall, F1 Score. The metrics were calculated using the annotated data from the recorded video dataset and the COCO-Pose dataset. To comprehensively evaluate the performance of our proposed gait analysis system, we utilized several key metrics. Each metric provides a different perspective on the system's accuracy and effectiveness. The metrics used include:

Precision: Precision indicates the ratio of correctly predicted positive observations to the total predicted positives.

$$\text{Precision } = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{2}$$

Recall: Recall measures the ratio of correctly predicted positive observations to all actual positives.

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{3}$$

F1 Score: The F1 Score is the harmonic mean of precision and recall, providing a single metric to balance the two.

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

Area Under the Curve (AUC): AUC measures the model's performance by plotting the true positive rate (TPR) against the false positive rate (FPR).

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR})\, d(\text{FPR}) \tag{5}$$

Mean Per Joint Position Error (MPJPE): MPJPE quantifies the average Euclidean distance between predicted and actual joint positions.

$$\text{MPJPE} = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{p}_i - \mathbf{g}_i\| \tag{6}$$

where $\mathbf{p}_i$ is the predicted position and $\mathbf{g}_i$ is the ground truth position of the i-th joint.

Percentage of Correct Keypoints at 10 pixels (PCK10): This metric calculates the percentage of keypoints correctly predicted within a 10-pixel radius of the ground truth.

$$\text{PCK}_{10} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\left(\|\mathbf{p}_i - \mathbf{g}_i\| < 10\right) \tag{7}$$

where $\mathbf{1}$ is an indicator function that returns 1 if the condition is true, otherwise 0.

These metrics provide a comprehensive evaluation of the performance of our proposed gait analysis system, ensuring that our gait analysis system is assessed from multiple performance angles, capturing accuracy, precision, error rates, and detection capabilities comprehensively. Besides, allowing us to assess its accuracy, precision, and recall in detecting and tracking gait patterns. By calculating these metrics on our dataset, we can quantify the system's performance and compare it to other state-of-the-art methods.

4.2.2 Data Analysis

In this study, we employed a markerless motion capture system to analyze the gait patterns of subjects walking on a more than 3 m walkway. The system utilized YOLO and ViTPose Standard Tool (Kinovea) to recognize keypoints and calculate knee, hip, and ankle flexion and extension angles. The subjects were allowed to start and end the gait cycles at their convenience, and three complete gait cycles were obtained per subject.

To define the gait cycle, it is essential to extract the target frame that accurately represents the spatiotemporal gait parameters of an individual. This can be achieved by tracking the person for a sufficient number of frames, allowing for the estimation of parameters such as the aspect ratio and the size of the closed area between the legs. In a comprehensive classification, the gait cycle is divided into two primary phases: stance and swing. The ratio of frames identified by the classifier for each primary phase can be calculated for each subject, providing a comprehensive understanding of the gait cycle. The natural ratio of primary gait phases is typically 40% for stance and 60% for swing in a complete gait cycle. By comparing the ratios of these primary phases, it is possible to assess the effectiveness of the gait cycle and identify any deviations from the natural ratio.

4.3 Mathematical Consideration

Detecting local extrema (both minima and maxima) in smoothed signals is a foundational technique in gait analysis, providing insights into the underlying patterns and behaviors of walking movements. By applying a uniform filter for smoothing followed by a systematic approach to identifying significant peaks and troughs, analysts can extract meaningful information from gait data with increased accuracy and reliability.

The point at index i to be a local minimum, it must satisfy the following conditions:

- $y_i < y_{i-1}$ (the value at index i is less than the value at the previous index).

- $y_i < y_{i+1}$ (the value at index i is less than the value at the next index).

For a point to be a local maximum, the conditions are reversed:

- $y_i > y_{i-1}$ (the value at index i is greater than the value at the previous index).

- $y_i > y_{i+1}$ (the value at index i is greater than the value at the next index).

We can represent these conditions using mathematical notation as follows:

For local minima:

$$\text{min} = \{i \mid y_i < y_{i-1} \text{ and } y_i < y_{i+1}\} \tag{8}$$

For local maxima:

$$\text{max} = \{i \mid y_i > y_{i-1} \text{ and } y_i > y_{i+1}\} \tag{9}$$

In the context of gait analysis, the unique patterns observed in the gait cycle can be used to identify individuals and distinguish between different gait patterns. By analyzing the height and width ratio of each sequence within a single gait cycle, distinctive curves are generated that serve as identifiers. The process identifies critical phases within a gait cycle based on the observation of three frames: the stance heel strike (initial contact), the middle frame heel off (mid-stance), and the last frame swing heel strike (terminal swing), corresponding to the maximum height and width in a gait cycle. The analysis of the aspect ratio curve, which plots the height-width ratio over time, enables the identification of a complete gait cycle through the detection of three local minima, indicating the key moments where the foot impacts the ground or lifts off, marking the start, mid-point, and end of a gait cycle (Hasan et al. 2022).

To programmatically identify and categorize these phases within the gait cycle, the following steps are taken:

Initial Contact (IC): The initial contact within a gait cycle is a critical phase that marks the moment the foot first touches the ground, starting a new cycle of movement. This point is essential in both clinical gait analysis and automated gait recognition systems, as it signifies the beginning of the stance phase of walking. Identifying the initial contact accurately is crucial for segmenting the gait cycle and analyzing the dynamics of walking patterns. In the provided context, the initial contact is determined by analyzing ankle distance data over a sequence of frames. The distance between ankles is plotted over time (or frame numbers), with the data color-coded to distinguish between positive (indicating the right leg is forward) and negative distances (indicating the left leg is forward). This color-coding helps visualize the alternation of steps and the transitions between the stance and swing phases for each leg.

Initial Contact Identification (ICI): The gait cycle frames are initially filtered to focus on points where the distance becomes positive, indicating the potential points for initial contact. By identifying the first positive distance within the gait cycle frames and adjusting the frame references accordingly, irrelevant data before the first confirmed initial contact is removed. Further refinement is done by adjusting the list of maximum points (mn_max) to ensure it aligns with the updated gait cycle frames, ensuring that the analysis focuses on complete gait cycles starting from an initial contact point.

In the field of human biomechanics, the gait cycle is a fundamental sequence governing human walking, and understanding its dynamics is crucial for enhancing rehabilitation strategies, optimizing prosthetic designs, and deepening our comprehension of human movement patterns. The gait cycle can be divided into eight distinct mathematical expressions, each representing pivotal phases of the gait cycle: For the calculation of gait cycles for all phases, mn_min is represented in Equation 9 and mn_mx is represented in Equation 10. Where, the point at index i represent the local minimum and maximum.

For mid-stance (MS) and initial swing (IS), the local maxima are not providing the accurate frame. However, if we subtract 4 from the local maxima, the output becomes more accurate. Where, 4 is the threshold value used to identify the accurate frame.

Initial Contact (IC): Represents the moment when the foot first makes contact with the ground during the gait cycle. This phase marks the beginning of the gait cycle, where the foot first makes contact with the ground.

$$IC = \mathrm{mn\_min}[i] \tag{10}$$

Loading Response (LR): As the name suggests, this phase denotes the moment when the heel of the foot strikes the ground. Refers to the phase where the body weight is transferred to the stance leg after initial contact.

$$LR = \frac{IC + \mathrm{mn\_max}[i]}{2} \tag{11}$$

Mid-Stance (MS): Occurs when the body weight is directly over the stance leg, and the leg is supporting the load.

$$MS = mn\_\mathrm{max}[i] - 4 \tag{12}$$

Terminal Stance (TS): This phase represents the final portion of the gait cycle, where the foot is in contact with the ground and the body's weight is supported by the foot. Marks the end of the stance phase, just before the swing phase begins.

$$TS = \mathrm{mn}_{-}\,\mathrm{min}[i+1] \tag{13}$$

Pre-Swing (PS): The transition phase between stance and swing, preparing for the leg to lift off.

$$PS = \frac{TS + \mathrm{mn\_max}[i+1]}{2} \tag{14}$$

Initial Swing (IS): The leg starts to swing forward, lifting off the ground.

$$IS = \mathrm{mn\_max}[i+1] - 4 \tag{15}$$

Mid-Swing (MS): This phase represents the middle portion of the gait cycle, where the foot is lifted off the ground and swung forward. The midpoint of the swing phase, is where the leg is at its highest point.

$$MSw = \frac{IS + mn\_min[i+2]}{2} \tag{16}$$

Terminal Swing (TS): This phase denotes the final portion of the gait cycle, where the foot is lifted off the ground and swung forward in preparation for the next step. The final part of the swing phase is before the next cycle begins.

$$TSw = \text{mn\_min}\,[i+2] \tag{17}$$

In the domain of human biomechanics, a set of eight mathematical expressions elucidates the essential phases constituting the gait cycle, which orchestrates the intricate process of human ambulation. These equations meticulously detail crucial events, from the inception of ground contact marked as "Initial Contact" to the fluid leg elevation observed during "Terminal Swing." Each formula serves as a blueprint for comprehending the biomechanical intricacies governing locomotion dynamics. Both researchers and practitioners leverage these insights to refine rehabilitation methodologies, optimize the design of prosthetic devices, and deepen our understanding of human movement patterns. This scholarly pursuit not only advances scientific knowledge but also fosters innovations that positively impact clinical practice and enhance the quality of life for individuals with mobility impairments.

4.4 Computation Cost

The computation cost analysis reveals the time required for various processing stages in our proposed approach. Our experimental video, with a duration of 9.0 seconds and a frame sampling rate of 1, consists of 280 frames. These frames were systematically processed through several stages of our methodology. The processing requirements are influenced by the video's length, resolution, aspect ratio, video quality, hardware performance, and other pertinent factors necessitating adjustments based on these parameters. The converted from video to frame processing time stands at a minimal 16.4 seconds, demonstrating the efficiency of our initial setup. YOLO and depth estimation processes take 2 minutes and 26.37 seconds, and 2 minutes and 33.4 seconds, respectively, highlighting the time-intensive nature of these tasks. STAF processing, necessary for accurate pose tracking, requires 3 minutes and 19.4 seconds, while CSRT and Lucas-Kanade optical flow processing times are significantly lower at 35.07 seconds and 19.58 seconds, respectively. ViTPose, despite its high accuracy, has a substantial processing time of 25 minutes and 37.18 seconds.

In total, the processing time sums up to 2107.40 seconds, resulting in an average processing time of 7.53 seconds per frame. This detailed breakdown emphasizes the computational efficiency and balance achieved across various stages of our method. The relatively high processing time of ViTPose is offset by the faster processes in other stages, ensuring that the overall performance remains practical for real-time applications. By optimizing these processing stages, our approach achieves high accuracy and efficiency, making it a robust solution for complex pose estimation tasks.

5. Results

This research study focuses on analyzing clinical gait patterns by examining human gait locomotion and the complex movements observed throughout the gait cycle. During the gait cycle analysis, the variation in ankle distance as individuals walk is observed, providing insights into their walking dynamics. Tracking ankle movement over time offers valuable information for rehabilitation strategies, prosthetic design, and enhancing our understanding of human locomotion. This research delves into clinical gait patterns by examining human gait locomotion and the intricate movements throughout the gait cycle. The focus on ankle distance dynamics, as illustrated in Figure 10, showcases the changes in ankle abduction during different walking phases, aiming to pinpoint key elements within the gait cycle, notably the initial contact phase.

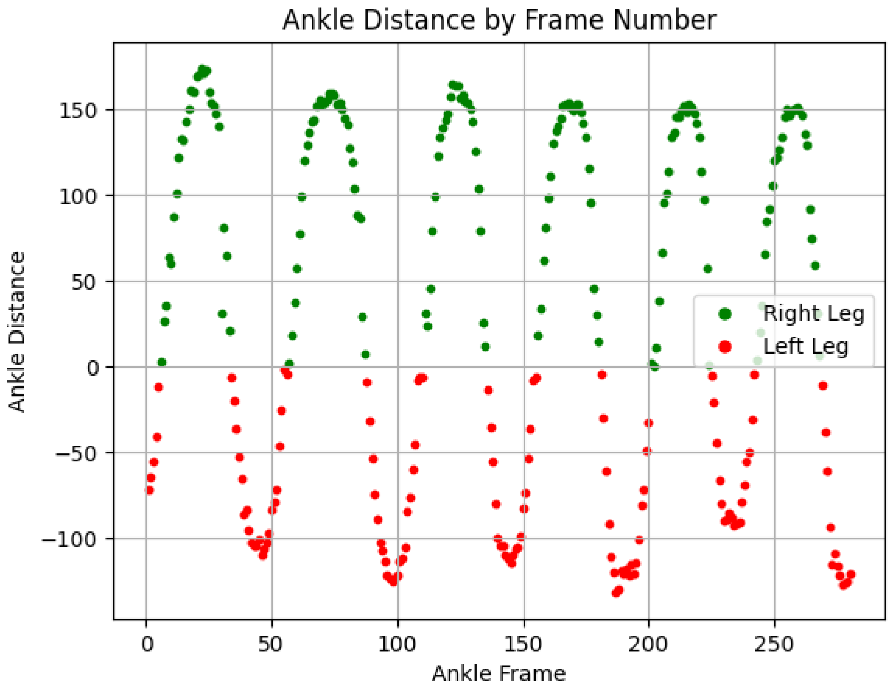

Figure 10. Ankle Distance Analysis During Gait Cycle

After identifying the first initial contact of a complete gait cycle, the study utilizes aspect ratio curves, as shown in Figure 11. While these curves provide valuable insights, accurately mapping the gait cycle presents challenges.

To address this, the researchers apply signal smoothing techniques, as evidenced by the blue curve in Figure 12. The comparison between the blue image and the reconstructed

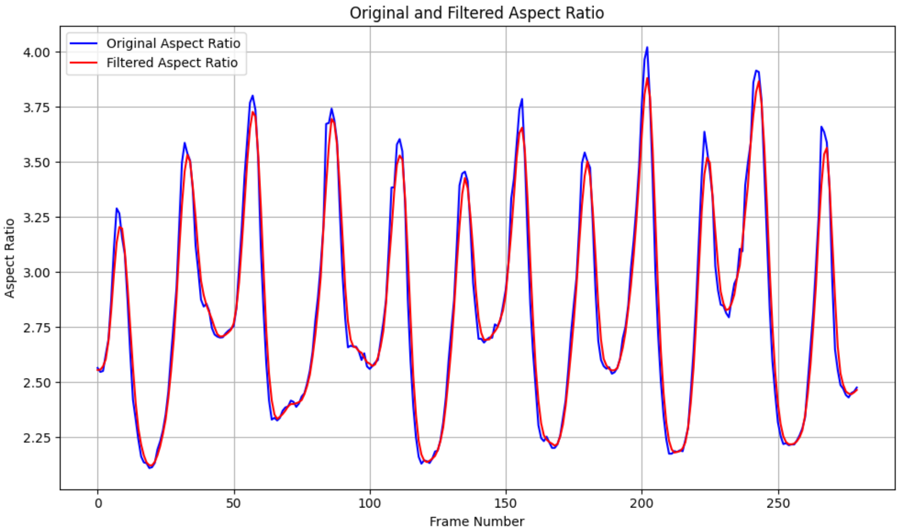

Figure 11. Comparative Analysis of Original and Filtered Aspect Ratio Curves

mesh images represented by the red curve proves that the proposed method correctly reconstructs the model.

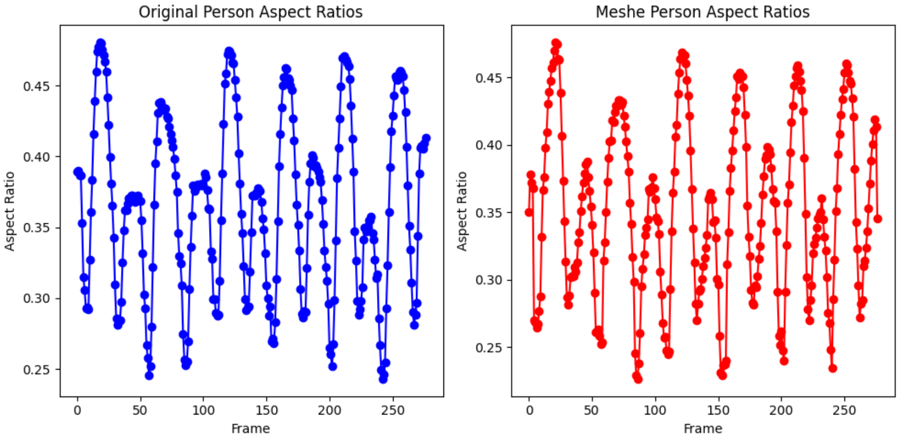

Figure 12. Comparative Aspect Ratios: Original and Meshe Person

Finally, we present the results of our analysis in the form of the complete gait cycle, broken down into its eight phases. This information is crucial for understanding the subjects' movements and the potential benefits for rehabilitation and medical applications.

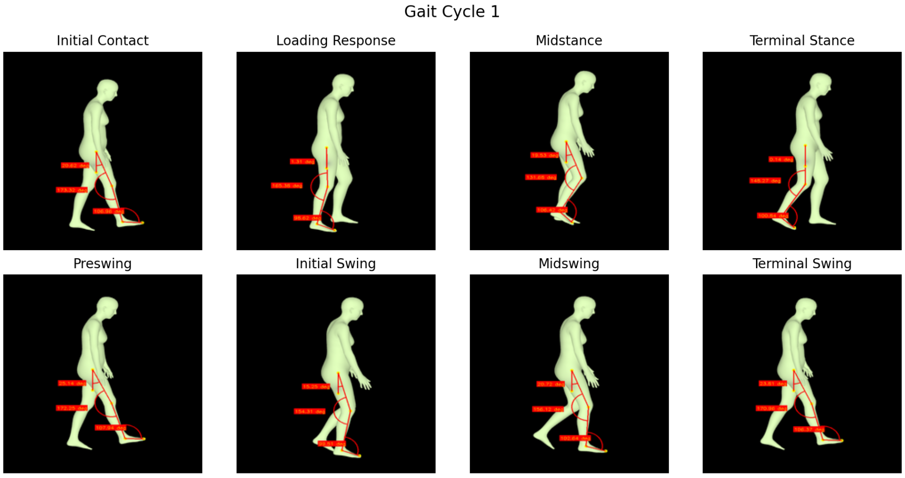

Figure 13. Gait Cycle Phases 1: Analysis and Insights

Furthermore, the study conducts a comprehensive examination of the gait cycle, dividing it into eight distinct phases. Figures 13 through 16 present detailed analyses and insights at each stage, revealing the complexities of human movement patterns at critical junctures in the cycle. Analyzing hip, knee, and ankle joint angles across different phases of walking, as shown in Figure 17, is significant for rehabilitation strategies, prosthetic innovations, and a comprehensive understanding of locomotor biomechanics.

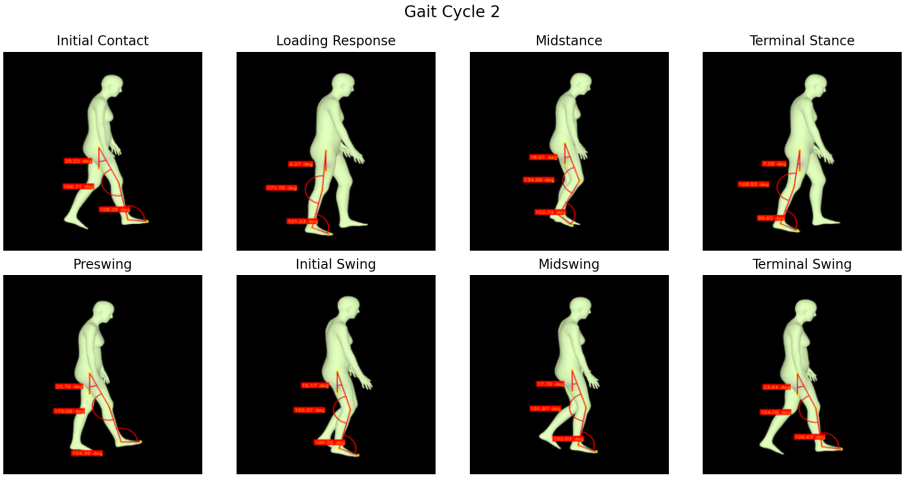

Figure 14. Gait Cycle Phases 2: Analysis and Insights

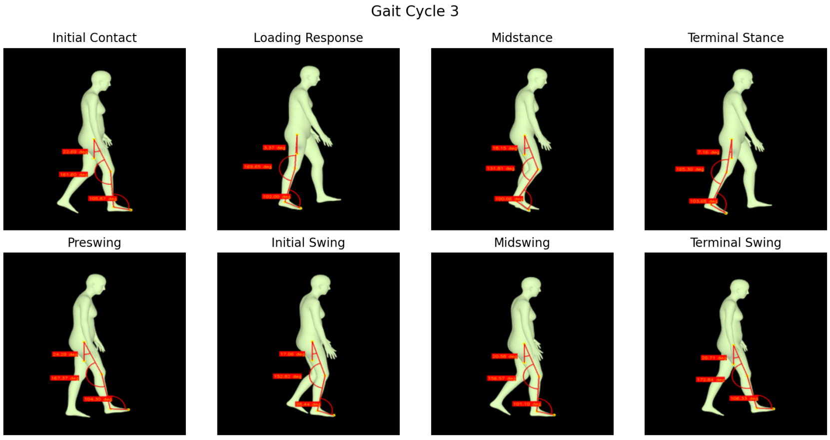

Figure 15. Gait Cycle Phases 3: Analysis and Insights

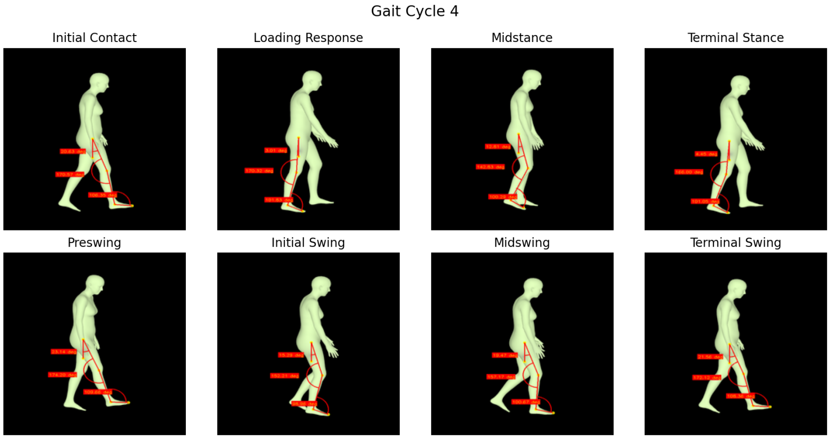

Figure 16. Gait Cycle Phases 4: Analysis and Insights

Figure 18 illustrates the joint angle analysis during different phases of the gait cycle. By examining how angles change at the hip, knee, and ankle throughout walking, we gain valuable insights into human locomotion. These findings can inform rehabilitation strategies, prosthetic design, and enhance our understanding of movement patterns.

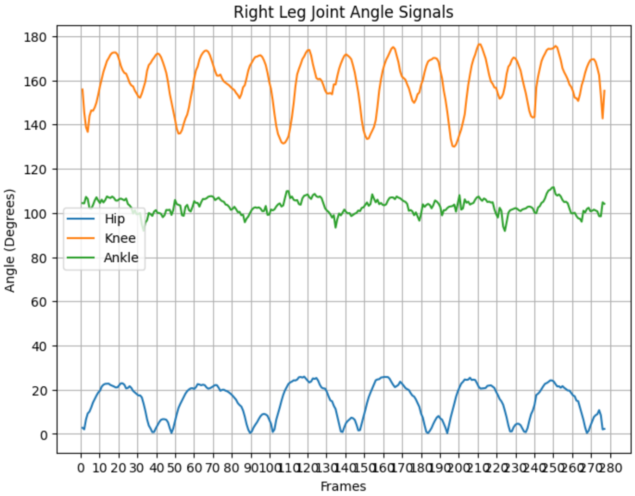

Figure 17. Frame Wise Joint Angle Variations Across Gait Cycle

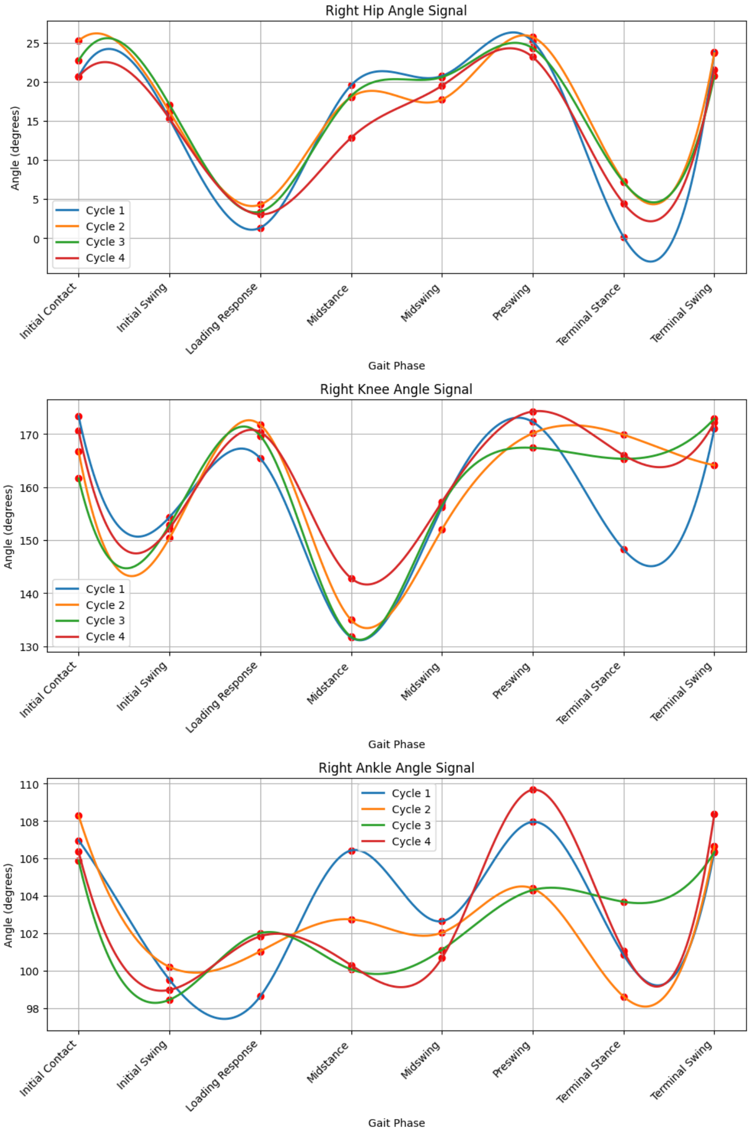

Figure 18. Joint Angle Analysis of Gait Cycle Phases

5.1 Result Comparison

This section represents a comparative analysis of joint angle annotation methods, specifically focusing on the hip, knee, and ankle joints. The Figure 19 showcases an annotated image of the angle of hip, knee, and ankle joint keypoints, calculated using of four different models: the Standard Tool (Ground Truth), YOLO, ViTPose, and the Proposed 3D ViTPose model. The Standard Tool provides a conventional baseline for angle measurement. YOLO demonstrates real-time object detection capabilities, marking key points with moderate accuracy. ViTPose employs vision transformers, offering enhanced precision in joint annotation. Finally, the Proposed 3D ViTPose model integrates advanced 3D pose estimation techniques, resulting in the most accurate and detailed representation of joint angles. This comparison highlights the advancements in joint angle calculation, emphasizing the superior performance of the proposed 3D ViTPose model.

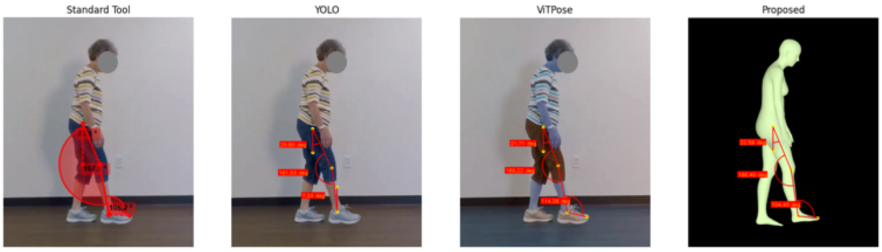

Figure 19. Annotated Frame of Standard Tool, YOLO, ViTPose, and Proposed 3D ViTPose with Angle

The Figure 20 illustrates a radar chart and Figure 21 bar chart comparing joint angles measured using four different tools: Standard Tool, YOLO, ViTPose, and the Proposed Model. The angles are plotted around the chart's perimeter, with each method's results forming a distinct polygon. The Standard Tool is represented in blue, YOLO in black, ViTPose in red, and the Proposed Model in pink. This visualization helps in evaluating and contrasting the accuracy and consistency of the joint angle measurements obtained by these methods. We can see from the Figure 20 that the triangles of proposed model and standard tool are more closed than the others. And Figure 21 shows that the proposed model's and Standard tool bar heights are more closed than the others.

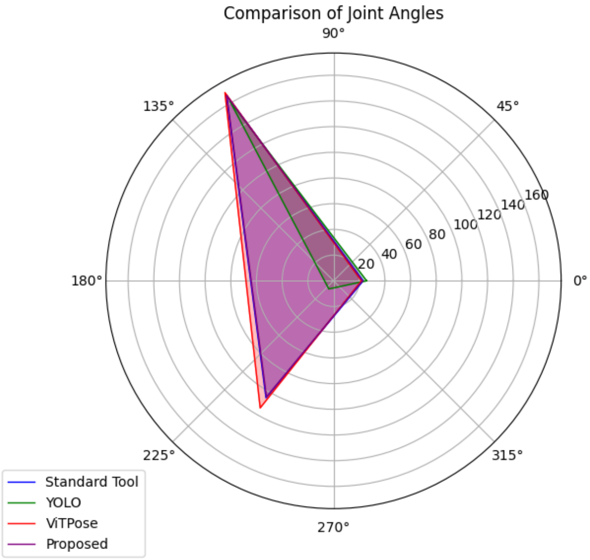

Figure 20. Comparison of joint angles for the Standard Tool, YOLO, ViTPose, and Proposed Model

We present the accuracy comparison of the proposed methodology to the standard tool, YOLO and the ViTPose model. The results show that the proposed methodology achieved an overall accuracy of 98.89%, with hip, knee, and ankle angle measurements being 97.81%, 99.58%, and 99.29% accurate, respectively, compared to the standard tool. Compared to the ViTPose model, the proposed methodology demonstrated significant improvements in accuracy, with overall accuracy of 98.89% compared to 94.70% for the ViTPose model. The hip, knee, and ankle angle measurements were 97.81%, 99.58%, and 99.29% accurate, respectively, compared to 93.99%, 98.73%, and 91.37% accurate for the ViTPose model.

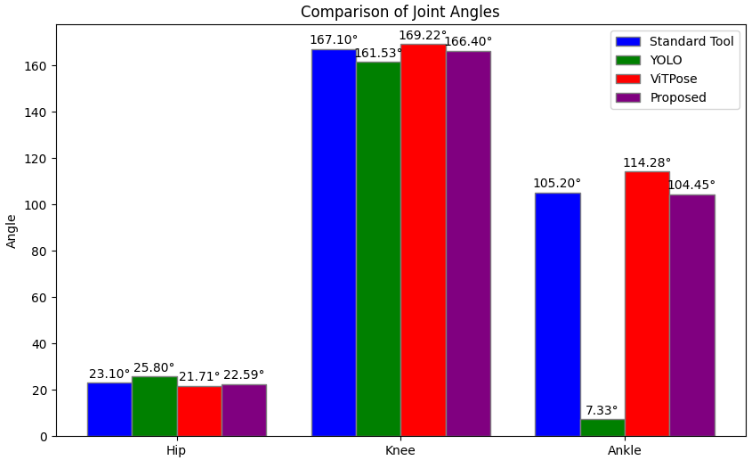

Figure 21. Comparison of Hip, Knee, and Ankle Angles for Different Models

These results demonstrate the superior performance of the proposed methodology in terms

Table 1. Classification Report of Proposed Model

Classification Report	**YOLOv8**	**YiTPose**	**Proposed Method**
Precision	0.64	0.92	0.98
Recall	0.64	0.98	100
F1 Score	0.64	0.95	0.99
AUC	0.44	0.48	0.50
MPJPE	3.80	0.20	0.12
PCK10	0.90	0.98	0.99

of accuracy, particularly for the ankle angle measurements, which showed a significant improvement over both the standard tool and the ViTPose model. The high accuracy of the proposed methodology has the potential to enhance rehabilitation strategies, optimize prosthetic designs, and deepen our understanding of human movement patterns.

The Table 1 presented offers a detailed classification report comparing the performance metrics of three different methods: YOLOv8, ViTPose, and the Proposed Method. The evaluation metrics included in this report are Precision, Recall, F1 Score, AUC, MPJPE, and PCK10.

The Proposed Method demonstrates superior performance across most metrics. It achieves the highest Precision at 0.98, the highest Recall at 100, and the highest F1 Score at 0.99. Additionally, it shows the best performance in terms of AUC (0.50), MPJPE (0.12), and PCK10 (0.99). In contrast, YOLOv8 and ViTPose show lower performance, with YOLOv8 having significantly lower metrics overall and ViTPose performing better than YOLOv8 but not surpassing the Proposed Method. This table underscores the effectiveness of the Proposed Method in gait analysis, showcasing its robustness and accuracy in comparison to the other evaluated methods.

This Table 2 shows the angle values of various models, so we can compare hip by hip, knee by knee, and ankle by ankle for different models. The measured hip angles are 23.10° (Standard Tool), 25.80° (YOLO), 21.17° (ViTPose), and 22.59° (Proposed Model). The measured knee angles are 167.10° (Standard Tool), 169.27° (YOLO), 61.53° (ViTPose), and 126.40° (Proposed Model). The measured ankle angles are 105.20° (Standard Tool), 7.33° (YOLO), 114.28° (ViTPose), and 96.45° (Proposed Model).

Table 2. Comparison of Hip, Knee, and Ankle Angles for Different Models

	Standard Tool	YOLO	ViTPose	Proposed
Hip	23.10	25.80	21.71	22.59
Knee	167.10	161.53	169.22	166.40
Ankle	105.20	7.33	114.28	104.45

Table 3. Joint Angle Accuracy of the Proposed Approach and Other Methods

Method	Accuracy (%)
YOLO	63.99
ViTPose	94.70
Proposed Approach	**98.89**
Standard Tool	100%

The table highlights (Table 3) the joint angle accuracy of different methods, showcasing our proposed approach's superior performance. YOLO's 63.99% accuracy stems from its primary focus on real-time object detection, which compromises its precision in detailed pose estimation tasks. ViTPose, though more accurate at 94.70%, suffers from high computational demands and occasional generalization issues across diverse datasets. Our proposed method addresses these limitations by integrating advanced algorithmic techniques that enhance precision without compromising on computational efficiency, ensuring robust performance in various real-time applications.

Our approach's 98.89% accuracy demonstrates its efficacy in overcoming the shortcomings of YOLO and ViTPose. By optimizing joint angle detection algorithms, we achieve high precision comparable to the standard tool (Kinovia) which is our ground truth values while maintaining computational efficiency suitable for real-time use. This balance allows our method to excel in practical applications, providing reliable and accurate joint angle measurements crucial for tasks such as gait analysis. Our focus on robustness and adaptability ensures consistent performance across different environments, making our solution a significant improvement over existing methods.

Compared to the ViTPose model, the proposed methodology demonstrated significant improvements in accuracy, from Table 3 with an overall accuracy of 98.89 compared to 94.70 for the ViTPose model. The hip, knee, and ankle angle measurements were 97.81, 99.58, and 99.29 accurate, respectively, compared to 93.99, 98.73, and 91.37 accurate for the ViTPose model. These results demonstrate the superior performance of the proposed methodology in terms of accuracy, particularly for the ankle angle measurements, which showed a significant improvement over both the standard tool and the ViTPose model. The high accuracy of the proposed methodology has the potential to enhance rehabilitation strategies, optimize prosthetic designs, and deepen our understanding of human movement patterns.

Finally the results demonstrate a significant enhancement in joint point accuracy through our methodology, which integrates 2D to 3D conversion techniques and ViTPose pose estimation. Furthermore, we have advanced the methodologies and algorithms for gait cycle calculation and phase identification by employing smooth signaling and local minima and maxima analysis. These improvements enable physiotherapists to achieve more precise assessments, facilitating better-informed decisions for treatment prescriptions.

6. Conclusions

In this study, we present a novel computer vision-based approach to human gait analysis, offering a robust, cost-effective, markerless, and user-friendly gait analysis system. Our proposed system addresses key limitations in current methodologies. Leveraging advanced technologies such as YOLO, ViTPose, and the Standard Tool (Kinovia), our system demonstrates significant potential for further enhancement in robustness and accuracy. Future developments may focus on error reduction through refined post-processing techniques and the establishment of standardized gait measurement protocols to elevate the system's performance and reliability. We progress to person detection and framing, which is the process of locating and recognising people inside video frames. For the next steps in pose estimation and gait analysis, this phase is essential. Next, we utilized YOLOv8, a state-of-the-art object detection algorithm, for pose estimation.

Depth estimation and 3D reconstruction further enhanced our understanding of human motion by providing depth information and reconstructing the 3D structure of the scene. These techniques enabled us to analyze gait patterns from different perspectives and angles. Temporal analysis, including motion analysis and video sequence analysis, allowed us to track changes in gait patterns over time. By analyzing video sequences and motion trajectories, we gained insights into how gait characteristics evolve dynamically.

Occlusion handling was another critical aspect addressed in our analysis. By resizing frames and implementing techniques to handle occlusions. We also explored ViTPose, a cutting-edge pose estimation model, capable of providing accurate joint predictions with 3D mesh outputs. ViTPose proved to be effective in both raw image and 3D mesh scenarios, offering high-fidelity representations of human poses.

Furthermore, we analyzed gait cycles and phases to understand the temporal dynamics of human motion. By segmenting gait cycles into distinct phases, such as initial contact, loading response, and swing, we gained insights into the sequence of events during the gait cycle. Frame-wise gait analysis and comparison of different methods, including Standard Tool, YOLO, ViTPose, and proposed 3D + ViTPose Angle, provided comprehensive evaluations of pose estimation accuracy and performance. We assessed the accuracy of each method relative to a standard tool and visualized the results using bar charts and polar plots.

In conclusion, developed a novel approach for analyzing human gait using markerless computer vision techniques. The proposed methodology includes a unique process for 2D to 3D reconstruction and pose estimation techniques to improve the accuracy of key joint points and calculating angles. Additionally, a new algorithm for calculating the gait cycle with phase has been created, which enhances the understanding of the overall structure of the subjects' movements. The proposed methodology has been compared to existing tools and models, demonstrating significant improvements in accuracy. The proposed methodology has the potential to have significant implications for the medical sector and rehabilitation practices, as it can enhance rehabilitation strategies, optimize prosthetic designs, and deepen our understanding of human movement patterns.

Medical Diagnostics, Osteopathic medicine, Comparative and Sports-related biomechanics, etc both benefit significantly from the research conducted in "Leveraging Markerless Computer Vision for Comprehensive Walking Automated Gait Analysis in Rehabilitation." Doctors can utilize the findings to enhance their understanding of gait patterns in patients undergoing rehabilitation, leading to more tailored treatment plans and better patient outcomes. On the other hand, biomedical scientists can leverage these insights to refine existing technologies and develop novel methodologies for automated gait analysis, thereby advancing the field of rehabilitation medicine and improving patient care.

6.1 Future Research Work

We have presented a novel approach for comprehensive walking automated gait analysis in rehabilitation using markerless computer vision techniques. Our methodology includes a unique process for 2D to 3D reconstruction and pose estimation to improve the accuracy of key joint points and angle calculations, as well as a new algorithm for calculating the gait cycle with phase. Despite the significant improvements in accuracy achieved by our proposed methodology, there are still several areas for future research. One potential direction for future work is to explore the use of additional deep learning models and architectures to further enhance the accuracy of gait analysis. This could involve investigating the use of convolutional neural networks (CNNs) or recurrent neural networks (RNNs) to better capture the temporal dynamics of gait movements. Another potential area for future research is to explore the use of additional sensors and modalities to enhance the accuracy of gait analysis. For example, incorporating data from wearable sensors or motion capture systems could provide additional insights into the biomechanics of gait and help to further refine the accuracy of our gait analysis methodology.

Finally, there is potential for the development of more advanced gait recognition systems that can be used in clinical settings to monitor and evaluate the progress of patients with various orthopedic and neurological conditions. This could involve the development of more sophisticated algorithms for gait phase segmentation and the integration of additional features, such as gait speed, cadence, and symmetry, to provide a more comprehensive assessment of gait function.

In conclusion, demonstrated the potential of marker-less computer vision techniques in gait analysis, offering a more accurate and efficient approach to analyzing human gait. The findings have significant implications for the medical sector and rehabilitation practices, as they can enhance rehabilitation strategies, optimize prosthetic designs, and improve overall patient outcomes. Future work in this area has the potential to further advance our understanding of gait function and provide more accurate and reliable gait analysis tools for clinical use.

6.2 Limitation of this Work

One of the major limitations of the proposed approach is the handling of motion in 3D frames. While the study has demonstrated improvements in gait analysis using markerless computer vision techniques, the accuracy of the results may be affected by the motion of the 3D frames. This is particularly relevant when analyzing complex movements or when dealing with large datasets.

Another limitation is the smooth 3D mesh generation process. Although the study has introduced a new algorithm for calculating the gait cycle with phase, the accuracy of the 3D mesh generation may be affected by the quality of the input data and the complexity of the gait cycles being analyzed.

To address these limitations, future work could focus on developing more robust algorithms for handling motion in 3D frames and improving the quality of 3D mesh generation. This could involve the use of more advanced deep learning models or the integration of additional sensors and modalities to provide more accurate and comprehensive gait analysis.

In conclusion, while this study has demonstrated the potential of markerless computer vision techniques in gait analysis, there are still limitations that need to be addressed to further enhance the accuracy and reliability of the results. Future work in this area could help to overcome these limitations and provide more advanced gait analysis tools for clinical use.

References

Ancans, Armands; Greitans, Modris; Cacurs, Ricards; Banga, Beate & Rozentals, Artis. 2021, Wearable sensor clothing for body movement measurement during physical activities in healthcare, *Sensors*, vol. 21, no. 6, p. 2068.

Arent, Ilja; Schmidt, Florian P; Botsch, Mario & Dürr, Volker. 2021, Marker-less motion capture of insect locomotion with deep neural networks pre-trained on synthetic videos, *Frontiers in behavioral neuroscience*, vol. 15, , p. 637806.

Avogaro, Andrea; Cunico, Federico; Rosenhahn, Bodo & Setti, Francesco. 2023, Markerless human pose estimation for biomedical applications: a survey, *arXiv preprint arXiv:2308.00519*.

Baker, Richard. 2006, Gait analysis methods in rehabilitation, *Journal of neuroengineering and rehabilitation*, vol. 3, , pp. 1–10.

Baker, Richard. 2007, The history of gait analysis before the advent of modern computers, *Gait & posture*, vol. 26, no. 3, pp. 331–342.

Bari, ASM Hossain & Gavrilova, Marina L. 2022, KinectGaitNet: Kinect-based gait recognition using deep convolutional neural network, *Sensors*, vol. 22, no. 7, p. 2631.

Bernal-Torres, Mario G; Medellín-Castillo, Hugo I & Arellano-González, Juan C. 2023, Development of a new low-cost computer vision system for human gait analysis: A case study, *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 237, no. 5, pp. 628–641.

Burnfield, M. 2010, Gait analysis: normal and pathological function, *Journal of Sports Science and Medicine*, vol. 9, no. 2, p. 353.

Cao, Zhe; Hidalgo, Gines; Simon, Tomas; Wei, Shih-En & Sheikh, Yaser. 2018, OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields, In: *arXiv preprint arXiv:1812.08008*.

Ceseracciu, Elena; Sawacha, Zimi & Cobelli, Claudio. 2014, Comparison of markerless and marker-based motion capture technologies through simultaneous data collection during gait: proof of concept, *PloS one*, vol. 9, no. 3, p. e87640.

Charmant, Joan. *Kinovea*. Available: https://www.kinovea.org/.

Cheng, Yang; Ugbolue, Ukadike C; Kerr, Andrew; Stankovic, Vladimir; Stankovic, Lina; Carse, Bruce; Kaliarntas, Konstantinos T & Rowe, Philip J. 2016, Autonomous Gait Event Detection with Portable Single-Camera Gait Kinematics Analysis System, *Journal of Sensors*, vol. 2016, .

Courtney, Jane & De Paor, Annraoi M. 2010, A monocular marker-free gait measurement system, *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 18, no. 4, pp. 453–460.

Davis III, Roy B; Ounpuu, Sylvia; Tyburski, Dennis & Gage, James R. 1991, A gait analysis data collection and reduction technique, *Human movement science*, vol. 10, no. 5, pp. 575–587.

Dunn, Marcus; Kennerley, Adam; Murrell-Smith, Zhane; Webster, Kate; Middleton, Kane & Wheat, Jon. 2023, Application of video frame interpolation to markerless, single-camera gait analysis, *Sports Engineering*, vol. 26, no. 1, p. 22.

Esquenazi, Alberto. 2014, Gait analysis in lower-limb amputation and prosthetic rehabilitation, *Physical Medicine and Rehabilitation Clinics*, vol. 25, no. 1, pp. 153–167.

Fernández-González, Pilar; Koutsou, Aikaterini; Cuesta-Gómez, Alicia; Carratalá-Tejada, María; Miangolarra-Page, Juan Carlos & Molina-Rueda, Francisco. 2020, Reliability of kinovea® software and agreement with a three-dimensional motion system for gait analysis in healthy subjects, *Sensors*, vol. 20, no. 11, p. 3154.

Gait, Mission. 2024, *Donate - Mission Gait*. Available: https://missiongait.org/donate/.

Geiger, Andreas; Lenz, Philip & Urtasun, Raquel. 2012, Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite, In: *Conference on Computer Vision and Pattern Recognition (CVPR)*.

Godard, Clément; Mac Aodha, Oisin; Firman, Michael & Brostow, Gabriel J. 2019, Digging into self-supervised monocular depth estimation, In: *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3828–3838.

Hasan, Md Mehedi; Ilham, Md Fatin; Sabbir, Khalid Saifullah & Aman, Nakib. 2022, Silhouette-based gait recognition using CNN-RNN, In: *2022 International Conference on Recent Progresses in Science, Engineering and Technology (ICRPSET)*, IEEE, pp. 1–4.

Heikel, Edvard & Espinosa-Leal, Leonardo. 2022, Indoor scene recognition via object detection and TF-IDF, *Journal of Imaging*, vol. 8, no. 8, p. 209.

Hellstén, Thomas; Karlsson, Jonny; Häggblom, Christian & Kettunen, Jyrki. 2022, Towards Accurate Computer Vision-Based Marker Less Human Joint Localization for Rehabilitation Purposes, In: *Scandinavian Conference on Health Informatics*, pp. 21–25.

Hellstén, Thomas; Karlsson, Jonny; Shamsuzzaman, Muhammed & Pulkkis, Göran. 2021, The potential of computer vision-based marker-less human motion analysis for rehabilitation, *Rehabilitation Process and Outcome*, vol. 10, , p. 11795727211022330.

Hii, Chang Soon Tony; Gan, Kok Beng; Zainal, Nasharuddin; Mohamed Ibrahim, Norlinah; Azmin, Shahrul; Mat Desa, Siti Hajar; van de Warrenburg, Bart & You, Huay Woon. 2023, Automated Gait Analysis Based on a Marker-Free Pose Estimation Model, *Sensors*, vol. 23, no. 14, p. 6489.

Hisham, Nur Amirah Hanisah; Ab Patar, Mohd Nor Azmi; Le, Chi Hieu; Mahmud, Jamaluddin; Lee, Hokyoo; Yamamoto, Shinichirou; Herawati, Lilik et al.. 2022, Walking gait analysis: kinovea versus motion capture system, In: *2022 IEEE 13th Control and System Graduate Research Colloquium (ICSGRC)*, IEEE, pp. 187–191.

Jacquelin Perry, M. 2010, Gait analysis: normal and pathological function, *New Jersey: SLACK*.

Jahangir, Faiza; Khan, Muhammad Attique; Alhaisoni, Majed; Alqahtani, Abdullah; Alsubai, Shtwai; Sha, Mohemmed; Al Hejaili, Abdullah & Cha, Jae-hyuk. 2023, A fusion-assisted multi-stream deep learning and ESO-controlled newton–raphson-based feature selection approach for human gait recognition, *Sensors*, vol. 23, no. 5, p. 2754.

Khan, Awais; Khan, Muhammad Attique; Javed, Muhammad Younus; Alhaisoni, Majed; Tariq, Usman; Kadry, Seifedine; Choi, Jung-In & Nam, Yunyoung. 2022, Human Gait Recognition Using Deep Learning and Improved Ant Colony Optimization., *Computers, Materials & Continua*, vol. 70, no. 2.

Kocabas, Muhammed; Athanasiou, Nikos & Black, Michael J. 2020, Vibe: Video inference for human body pose and shape estimation, In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5253–5263.

Konz, Latisha; Hill, Andrew & Banaei-Kashani, Farnoush. 2022, ST-DeepGait: A spatiotemporal deep learning model for human gait recognition, *Sensors*, vol. 22, no. 20, p. 8075.

Kumar, Sandeep; Gopinath, Karthik; Rocchi, Laura; Sukumar, Poorna Talkad; Kulkarni, Suyameendra & Sampath, Jayanth. 2018, Towards a portable human gait analysis & monitoring system, In: *2018 International Conference on Signals and Systems (ICSigSys)*, IEEE, pp. 174–180.

Lam, Winnie WT; Tang, Yuk Ming & Fong, Kenneth NK. 2023, A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation, *Journal of NeuroEngineering and Rehabilitation*, vol. 20, no. 1, p. 57.

Liang, Shengyun; Zhang, Yu; Diao, Yanan; Li, Guanglin & Zhao, Guoru. 2022, The reliability and validity of gait analysis system using 3D markerless pose estimation algorithms, *Frontiers in Bioengineering and Biotechnology*, vol. 10, , p. 857975.

Lin, Tsung-Yi; Maire, Michael; Belongie, Serge; Hays, James; Perona, Pietro; Ramanan, Deva; Dollár, Piotr & Zitnick, C Lawrence. 2014, Microsoft coco: Common objects in context, In: *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, Springer, pp. 740–755.

Majerník, Jaroslav. 2015, Reconstruction of Human Motion Trajectories to Support Human Gait Analysis in Free Moving Subjects, *Computational Intelligence, Medicine and Biology: Selected Links*, pp. 57–77.

Moran, Matthew F; Rogler, Isabella C & Wager, Justin C. 2023, Inter-session repeatability of marker-less motion capture of treadmill running gait, *Applied Sciences*, vol. 13, no. 3, p. 1702.

Moro, Matteo; Marchesi, Giorgia; Hesse, Filip; Odone, Francesca & Casadio, Maura. 2022, Markerless vs. marker-based gait analysis: A proof of concept study, *Sensors*, vol. 22, no. 5, p. 2011.

Muro-De-La-Herran, Alvaro; Garcia-Zapirain, Begonya & Mendez-Zorrilla, Amaia. 2014, Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications, *Sensors*, vol. 14, no. 2, pp. 3362–3394.

Park, Suah; Bae, Byunghoon; Kang, Kyungmin; Kim, Hyunjee; Nam, Mi Song; Um, Jumyung & Heo, Yun Jung. 2023, A deep-learning approach for identifying a drunk person using gait recognition, *Applied Sciences*, vol. 13, no. 3, p. 1390.

Sharif, Muhammad Imran; Khan, Muhammad Attique; Alqahtani, Abdullah; Nazir, Muhammad; Alsubai, Shtwai; Binbusayyis, Adel & Damaševičius, Robertas. 2022, Deep learning and kurtosis-controlled, entropy-based framework for human gait recognition using video sequences, *Electronics*, vol. 11, no. 3, p. 334.

Shi, Ling-Feng; Liu, Zhong-Ye; Zhou, Ke-Jun; Shi, Yifan & Jing, Xiao. 2023, Novel deep learning network for gait recognition using multimodal inertial sensors, *Sensors*, vol. 23, no. 2, p. 849.

Talib, Moahaimen; Al-Noori, Ahmed HY & Suad, Jameelah. 2024, YOLOv8-CAB: Improved YOLOv8 for Real-time object detection, *Karbala International Journal of Modern Science*, vol. 10, no. 1, p. 5.

Tao, Weijun; Liu, Tao; Zheng, Rencheng & Feng, Hutian. 2012, Gait analysis using wearable sensors, *Sensors*, vol. 12, no. 2, pp. 2255–2283.

Tunca, Can; Pehlivan, Nezihe; Ak, Nağme; Arnrich, Bert; Salur, Gülüstü & Ersoy, Cem. 2017, Inertial sensor-based robust gait analysis in non-hospital settings for neurological disorders, *Sensors*, vol. 17, no. 4, p. 825.

Ultralytics. 2024, *COCO*. Available: https://docs.ultralytics.com/datasets/pose/coco/.

Whittle, Michael W. 1996, Clinical gait analysis: A review, *Human movement science*, vol. 15, no. 3, pp. 369–387.

Whittle, Michael W. 2014, *Gait analysis: an introduction*, Butterworth-Heinemann.

Wikipedia. 2024, *Lucas–Kanade method*. Available: https://en.wikipedia.org/wiki/Lucas%E2%80%93Kanade_method.

Xu, Yufei; Zhang, Jing; Zhang, Qiming & Tao, Dacheng. 2022, ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation, In: *Advances in Neural Information Processing Systems*.

Zidani, Ghania; DJARAH, Djalal; BENMAKHLOUF, Abdslam & KHETTACHE, Laid. 2024, OPTIMIZING PEDESTRIAN TRACKING FOR ROBUST PERCEPTION WITH YOLOv8 AND DEEPSORT, *Applied Computer Science*, vol. 20, no. 1, pp. 72–84.

Appendix A

VIBE: VIBE, or Visual Inertial Body Estimation, is a sophisticated technology that merges visual information from cameras with inertial data to accurately estimate body pose and movement. In the realm of gait analysis, VIBE holds significant importance as it enables precise tracking of body joint movements during walking or running activities. By combining visual and inertial data, VIBE enhances the accuracy of gait analysis, providing detailed insights into biomechanical patterns and abnormalities. This technology facilitates real-time monitoring of gait dynamics, allowing for immediate feedback and adjustment of rehabilitation strategies. Moreover, VIBE's versatility and adaptability to different environments make it invaluable for clinical assessments and personalized rehabilitation plans. In my work, the integration of VIBE enhances the precision of gait analysis, enabling comprehensive evaluations and informed decision-making in rehabilitation processes.

CSRT: CSRT, or Continuously Adaptive Mean Shift with Kernel Correlation Filter Tracking, is a robust tracking algorithm utilized in computer vision for gait analysis. It effectively tracks and follows specific features or objects of interest within a video sequence, allowing for the accurate monitoring of human gait patterns over time. CSRT employs adaptive techniques to account for variations in appearance, scale, and orientation, ensuring reliable tracking even in challenging conditions such as occlusions or changes in lighting. In my work, the integration of CSRT enhances the precision and efficiency of gait analysis by enabling the continuous and reliable tracking of key anatomical landmarks or body segments during movement. This facilitates comprehensive assessments of gait dynamics and abnormalities, contributing to informed decision-making in rehabilitation strategies and clinical interventions.

Lucas-Kanade: The Lucas-Kanade method is a widely used technique in computer vision for optical flow estimation, developed by Bruce D. Lucas and Takeo Kanade. It is particularly valuable in gait analysis due to its ability to track motion by comparing intensity gradients between consecutive frames of video. This method calculates the flow of pixels between frames, allowing for the measurement of movements in human gait. In gait analysis research, the Lucas-Kanade model plays a crucial role in accurately tracking and analyzing the movement patterns of individuals during walking or running. Its importance lies in providing quantitative data on gait parameters such as stride length, step duration, and joint angles, aiding in the assessment of biomechanical characteristics and the evaluation of rehabilitation interventions. By leveraging the Lucas-Kanade method, researchers can enhance the precision and efficiency of gait analysis, leading to advancements in understanding human locomotion and optimizing rehabilitation strategies (Wikipedia 2024).

YOLO: YOLO, or You Only Look Once, is a state-of-the-art object detection algorithm used in computer vision for gait analysis. It excels in real-time detection of pedestrians and their key body features, making it invaluable for gait recognition and understanding. YOLO's ability to detect moving objects swiftly and accurately, even in complex scenes, enhances the efficiency of gait analysis by providing precise data on pedestrian movements. In my work, integrating YOLO enables comprehensive and automated gait analysis, facilitating the identification of gait abnormalities and aiding in the development of tailored interventions for rehabilitation and clinical assessment.

Standard Tool (Kinovia): Kinovea is a free video analysis software widely utilized in gait analysis, sports science, and training evaluation. It allows users to capture, slow down, compare, annotate, and measure motion in videos, making it valuable for studying walking patterns and biomechanics. Key features include the measurement of joint angles, stride length, and step duration, providing essential quantitative data for gait analysis. Kinovea's primary benefits are its user-friendly interface and the ability to perform analyses without physical sensors. Moreover, it is cost-effective and can be used for rehabilitation measurements.

Kinovea plays a pivotal role in gait analysis research by enabling detailed examination of gait parameters and movement patterns. Researchers can analyze video recordings of walking individuals to identify abnormalities, assess interventions, and monitor progress over time.

By leveraging Kinovea (Charmant), researchers can enhance the accuracy and efficiency of gait analysis, leading to a deeper understanding of human locomotion and improved rehabilitation strategies (Hisham et al. 2022). The aim of this thesis is to compare the accuracy of our proposed method for measuring the kinematic angles of the ankle, knee, and hip with Kinovea, focusing on the detection and analysis of movements for lower limb rehabilitation.

VICON: Vicon is a leading provider of motion capture systems renowned for their accuracy and reliability in gait analysis. Utilizing advanced camera technology and proprietary software, Vicon systems capture precise movement data by tracking reflective markers placed on the subject's body. This data allows researchers and clinicians to analyze various aspects of human gait, such as stride length, cadence, and joint angles, facilitating the assessment of biomechanical abnormalities and informing treatment strategies. Vicon's motion capture technology plays a crucial role in understanding human movement patterns, and enhancing research in fields like biomechanics, rehabilitation, and sports science.