

Review

Not peer-reviewed version

---

# Optical Sensor Based Approaches in Obesity Detection: A Literature Review of Gait Analysis, Pose Estimation, and Human Voxel Modeling

---

[Mohamed Moncef Ben Khelifa](#)\*, [Sabrine Dhaouadi](#), [Ala Balti](#), [Pascale Duche](#)

Posted Date: 19 May 2025

doi: 10.20944/preprints202505.1399.v1

Keywords: optical sensors; obesity detection; gait analysis; pose estimation; human voxel modeling



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

# Optical Sensor Based Approaches in Obesity Detection: A Literature Review of Gait Analysis, Pose Estimation, and Human Voxel Modeling

Mohamed Moncef BEN KHELIFA \*, Sabine DHAOUADI, Ala BALTI <sup>1</sup> and Pascale DUCHE

J-AP2S "LABORATOIRE JEUNESSE – ACTIVITÉ PHYSIQUE ET SPORTIVE – SANTÉ", Toulon University

\* Correspondence: khelifa@univ-tln.fr

## Highlights

- Comprehensive analysis of cutting-edge optical and vision-based systems for obesity detection through detailed gait pattern analysis.
- Evaluation of pose estimation algorithms, optical sensing technologies, and 3D voxel-based modeling approaches for quantitative obesity assessment.
- Critical examination of technical and implementation challenges for developing real-time, field-deployable, and AI-integrated obesity screening systems.
- Detailed review of how non-contact optical sensor methodologies enable quantitative assessment of obesity-related gait patterns and postural deviations.
- Comparative evaluation of markerless pose estimation frameworks (including OpenPose and MediaPipe) and their effectiveness in detecting obesity-related kinematic abnormalities.
- Technical assessment of infrared depth sensor applications in 3D voxel model reconstruction for contactless body composition analysis in obesity screening.
- Identification of current limitations in optical sensor approaches, including algorithm bias, environmental robustness constraints, and deployment scalability challenges.
- Analysis of emerging research directions in obesity detection, including edge-AI implementation, federated learning methodologies, and hybrid wearable-optical sensor fusion systems.

**Abstract:** Optical sensor technologies are reshaping obesity detection by enabling non-invasive, dynamic analysis of biomechanical and morphological biomarkers. This review synthesizes recent advances in three key areas: optical gait analysis, vision-based pose estimation, and depth-sensing voxel modeling. Gait analysis leverages optical sensor arrays and video systems to identify obesity-specific deviations, such as reduced stride length and asymmetric movement patterns. Pose estimation algorithms—including markerless frameworks like OpenPose and MediaPipe—track kinematic patterns indicative of postural imbalance and altered locomotor control. Human voxel modeling reconstructs 3D body composition metrics, such as waist-hip ratio, through infrared-depth sensing, offering precise, contactless anthropometry. Despite their potential, challenges persist in sensor robustness under uncontrolled environments, algorithmic biases in diverse populations, and scalability for widespread deployment in existing health workflows. Emerging solutions such as federated learning and edge computing aim to address these limitations by enabling multimodal data harmonization and portable, real-time analytics. Future priorities involve standardizing validation protocols to ensure reproducibility, optimizing cost-efficacy for scalable deployment, and integrating optical systems with wearable technologies for holistic health monitoring. By shifting obesity diagnostics from static metrics to dynamic, multidimensional profiling, optical sensing paves the way for scalable public health interventions and personalized care strategies.

**Keywords:** optical sensors; obesity detection; gait analysis; pose estimation; human voxel modeling

---

## 1. Introduction

### 1.1. Background and Motivation

Obesity has emerged as a critical global health concern, affecting individuals across all age groups and geographic regions. In 2022, more than 1 billion people worldwide—equivalent to one in eight—were living with obesity, with adult rates more than doubling and childhood and adolescent rates quadrupling since 1990 [1,2]. If current trends continue, projections indicate that over half the global population will be overweight or obese within the next 12 years, with the prevalence of obesity alone expected to rise from 14% to 24% by 2035, affecting nearly 2 billion people [2] [3]. Notably, the increase in obesity is steepest among children and adolescents: the percentage of boys affected is projected to double from 10% to 20% and girls from 8% to 18% between 2020 and 2035 [3]. The burden is not limited to high-income countries; rapid increases are also observed in low- and middle-income nations, particularly in Asia and Africa, where childhood overweight rates have surged by nearly 24% since 2000 [4].

Obesity is a multifactorial chronic disease with complex physiological, psychological, and socioeconomic implications. It is a major risk factor for non-communicable diseases, including cardiovascular disease, type 2 diabetes, osteoarthritis, and several cancers [1,2]. The economic impact is substantial, with annual global healthcare costs attributable to obesity exceeding \$2 trillion [3]. Despite a strong evidence base for effective interventions, implementation remains patchy, and the epidemic continues to escalate [1].

Early detection and intervention are paramount, as obesity often leads to progressive impairment in physical function, quality of life, and long-term health outcomes [1,2]. Traditional screening methods, such as body mass index (BMI), provide only a static snapshot and may not capture the early biomechanical changes linked to excess weight. Increasing evidence highlights the importance of functional markers—especially those related to movement and gait—as early indicators of obesity-related health risks [5,6].

### 1.2. Gait as a Diagnostic Tool

Human gait is a complex, dynamic process that reflects the integration of neuromuscular, skeletal, and metabolic systems. In the context of obesity, gait analysis has emerged as a valuable tool for identifying early biomechanical alterations that precede clinical symptoms [5]. Obese individuals—both adults and children—consistently demonstrate distinct gait characteristics: reduced stride length, slower walking speed, increased double support time, and greater asymmetry in joint loading, particularly at the hip, knee, and ankle [5]. These changes are not merely compensatory responses to increased body mass; they are also predictive of future musculoskeletal complications, reduced mobility, and diminished quality of life.

Recent studies using inertial measurement units (IMUs) and deep learning models have shown that gait patterns can accurately differentiate between normal-weight and obese adolescents, achieving classification accuracies as high as 97% [6]. Obese individuals exhibit shorter step lengths, slower speeds, and greater variability in gait, supporting the use of gait metrics as sensitive markers for early detection and monitoring of obesity-related functional decline [5,6]. Unlike static measures such as BMI or waist circumference, gait analysis provides a dynamic assessment of how excess weight affects daily movement and joint stress.

However, traditional gait analysis methods—such as marker-based motion capture systems and force plates—are often expensive, time-consuming, and limited to specialized laboratories. These constraints have historically restricted the use of gait analysis in routine clinical or community-based screening.

### 1.3. Shift in Technology: Toward Optical and Computational Sensing

Technological advances over the past decade have transformed the landscape of biomechanical assessment. Optical sensor systems, including RGB-D cameras (e.g., Microsoft Kinect, Intel RealSense), stereo vision, and monocular camera setups, now enable robust, markerless motion capture in real-world environments. These systems, when integrated with artificial intelligence (AI) and machine learning algorithms, allow for the extraction of detailed gait and posture metrics from simple video or depth data, making large-scale, non-invasive health screening feasible and cost-effective.

Markerless pose estimation frameworks—such as OpenPose, MediaPipe, and HRNet—can extract 2D or 3D skeletal keypoints from video input in real time, enabling efficient analysis of joint trajectories, angles, and coordination. These tools have been successfully applied to detect gait abnormalities in a range of clinical populations, including those with neurological and metabolic disorders, and are now being adapted for obesity screening. Additionally, 3D voxel modeling techniques derived from multi-view images or depth data provide volumetric insights into body composition, posture, and load distribution—factors highly relevant to obesity diagnosis and monitoring.

The integration of AI-powered analysis with optical sensing offers several advantages:

- **Non-invasiveness:** No physical contact or markers required, increasing user comfort and compliance.
- **Scalability:** Portable and low-cost systems enable deployment in diverse settings, from clinics to homes and schools.
- **Automation:** AI-driven pipelines facilitate rapid, objective assessment, reducing operator dependency and human error.
- **Personalization:** Continuous monitoring allows for individualized feedback and early intervention.

Despite these advances, challenges remain. There are ongoing debates regarding the reliability and validity of markerless optical systems compared to gold-standard laboratory instrumentation. Most algorithms are trained on normative datasets with limited representation of obese or morphologically diverse individuals, raising concerns about generalizability and algorithmic bias. Technical issues such as occlusion, clothing variability, and limited ground-truth data further complicate validation and deployment.

### 1.4. Scope and Objectives of the Review

Given these developments, the present review aims to consolidate and critically evaluate current optical sensor-based approaches for obesity detection, with a focus on three key domains:

- Optical gait analysis systems that derive spatiotemporal and kinematic metrics from video or depth data.
- Vision-based pose estimation frameworks that infer body mechanics from 2D/3D skeletal reconstructions.
- 3D voxel modeling techniques that provide volumetric insights into posture and body shape relevant to obesity diagnosis.

This review is intended for a multidisciplinary audience, including researchers and developers in biomedical sensing, artificial intelligence, health technology, biomechanics, and clinical diagnostics. By synthesizing findings from recent literature, we aim to:

- Provide an accessible overview of state-of-the-art methodologies and comparative system performance.

- Discuss validation, accessibility, and ethical considerations in deploying these technologies.
- Highlight both the current potential and limitations of optical sensor-based systems.
- Identify opportunities for future research and clinical translation.

In conclusion, the integration of gait analysis, pose estimation, and voxel modeling through optical sensing technologies holds transformative promise for early, individualized, and scalable obesity diagnostics. This is particularly significant for children and adolescents, where early detection and intervention can have lifelong health benefits. By bridging advances in sensing technology and AI with clinical needs, the field is poised to make a substantial impact on global efforts to curb the obesity epidemic.

## 2. Review Methodology

This literature review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines. While this article is not a formal systematic review, we adhered to PRISMA principles to ensure methodological rigor and reproducibility. The primary research question guiding this review was expanded to encompass multiple dimensions:

*“What are the current optical sensor technologies and methodological approaches used for detecting and analyzing obesity through gait analysis, pose estimation, and human voxel modeling?”*

Secondary questions include:

1. *How has the landscape of optical sensor technology for obesity detection evolved since 2000?*
2. *What are the comparative advantages of different optical sensing modalities for obesity assessment?*
3. *What methodological challenges exist in validating these technologies across diverse populations?*
4. *How do optical sensor approaches compare with traditional obesity assessment methods?*

We included studies published in English or French, prioritizing those with significant academic influence (e.g., citation frequency, high-impact venues) to ensure methodological robustness and relevance.

### 2.1. Search Strategy and Information Sources

As shown in Table 1, a comprehensive search was conducted using seven primary electronic databases to ensure wide coverage across medical, engineering, and computer science domains:

**Table 1.** The Primary databases used for the literature review study.

Database	Rationale for Inclusion	Field Coverage
PubMed/MEDLINE	Core biomedical literature	Medicine, biomechanics, clinical validation
Scopus	Broad multidisciplinary coverage	Engineering, computer science, healthcare
IEEE Xplore	Engineering and computing focus	Signal processing, sensor design, algorithms
ACM Digital Library	Computing research	Computer vision, machine learning
ScienceDirect	Multidisciplinary science platform	Optical engineering, biomechanics
Web of Science	Citation tracking capability	Cross-disciplinary research
Google Scholar	Grey literature and technical reports	Emerging technologies, pre-prints

The search period covered January 2006 through April 2025, capturing both foundational works and recent technological advances. Additionally, we employed citation tracking (both forward and backward) to identify seminal papers that may have been missing in the database searches.

**Example PubMed Search (English):**

("obesity" OR "overweight") OR  
 ("gait analysis" OR "stride" OR "walking pattern") AND  
 ("optical sensor" OR "OptoGait" OR "pose estimation" OR "OpenPose" OR "MediaPipe" OR  
 "Kinect" OR "voxel model") AND  
 ("validation" OR "accuracy" OR "biomechanics" OR "machine learning")

*2.2. Inclusion and Exclusion Criteria*

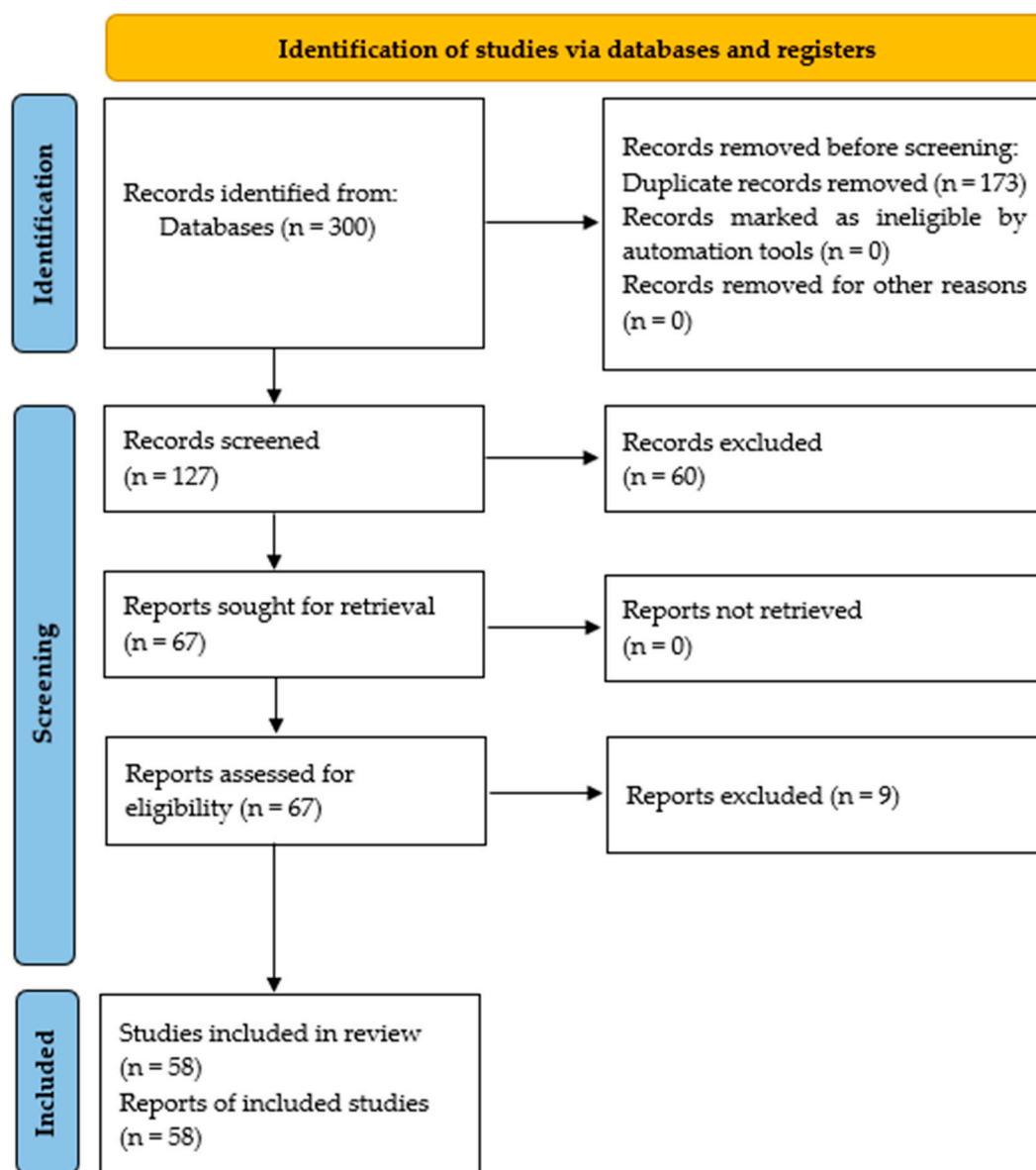
The eligibility criteria were refined and expanded from the original methodology to ensure precise inclusion of relevant studies. Table 2 presents the detailed inclusion and exclusion criteria.

**Table 2.** The inclusion & exclusion criterias.

Category	Inclusion Criteria	Exclusion Criteria
<b>Language</b>	English	Any other languages
<b>Publication Type</b>	Peer-reviewed full-text Conference and journal articles.  Thesis documents	Abstracts without full papers, editorials, opinion pieces
<b>Population / Sample (P/S)</b>	Human participants of any age group, with or without obesity	Studies involving only animals or synthetic (non-human) datasets
<b>Phenomenon / Intervention (PI/I)</b>	Use of optical sensing (e.g., OptoGait), pose estimation (e.g., OpenPose, MediaPipe), voxel modeling (e.g., Kinect) for gait or anthropometric analysis  Studies using inertial sensors for comparison	Studies using only wearable sensors, or manual observation with no imaging or optical component
<b>Design (D/C)</b>	Cross-sectional, observational, technical validation, mixed-methods, or experimental studies	Purely theoretical models without empirical validation, no performance evaluation
<b>Outcomes (O/E)</b>	Gait parameters (e.g., stride length, toe clearance, joint angles), obesity markers (e.g., body volume, asymmetry), diagnostic performance, usability, or real-world deployability	No relevant metrics related to gait, anthropometry, or obesity-specific detection
<b>Research Type (R)</b>	Quantitative, technical validation, mixed-methods studies	Literature reviews, non-empirical articles

### 2.3. Study Selection Process

The study selection process followed the PRISMA 2020 guidelines and is visually represented in Figure 1, which captures the flow of information through different phases of the review. A total of 300 records were retrieved. After removing duplicates, 127 titles and abstracts were screened. Of these, 67 full-text articles were assessed for eligibility. A final total of 58 articles met the inclusion criteria.



**Figure 1.** Selection of the relevant papers based on the PRISMA 2020 Flow Diagram.

### 2.4. Quality Assessment and Risk of Bias

A methodical quality assessment process was implemented to evaluate the included studies. Given the interdisciplinary nature of the research spanning engineering, computer science, and clinical domains, we developed a custom quality assessment tool that incorporates elements from:

- The Joanna Briggs Institute (JBI) Critical Appraisal Tools
- The Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2)

- Additional technical criteria specific to optical sensing technologies

Table 3 outlines the quality assessment criteria used to evaluate included studies.

**Table 3.** Quality assessment criteria used for evaluation in this work.

Domain	Assessment Criteria	Scoring
Study Design	<ul style="list-style-type: none"> <li>- Clear research objectives and questions</li> <li>- Appropriate study design for objectives</li> <li>- Adequate sample size with power analysis where appropriate</li> </ul>	0-3 points
Participant Selection	<ul style="list-style-type: none"> <li>- Clear inclusion/exclusion criteria</li> <li>- Representative sample of target population</li> <li>- Appropriate participant characteristics reported (age, sex, BMI, health status)</li> </ul>	0-3 points
Technical Methodology	<ul style="list-style-type: none"> <li>- Detailed description of hardware specifications</li> <li>- Comprehensive explanation of algorithms and processing pipelines</li> <li>- Appropriate calibration and validation procedures</li> <li>- Clearly defined parameters and metrics</li> </ul>	0-4 points
Reference Standard	<ul style="list-style-type: none"> <li>- Use of appropriate gold standard or reference measures</li> <li>- Proper implementation of reference measures</li> <li>- Blinding between test and reference standard where applicable</li> </ul>	0-3 points
Data Analysis	<ul style="list-style-type: none"> <li>- Appropriate statistical methods</li> <li>- Proper handling of missing data</li> <li>- Appropriate performance metrics reported (e.g., accuracy, precision, recall)</li> <li>- Consideration of confounding variables</li> </ul>	0-4 points
Results Reporting	<ul style="list-style-type: none"> <li>- Complete reporting of all planned outcomes</li> <li>- Appropriate presentation of results (tables, figures)</li> <li>- Comprehensive discussion of limitations</li> <li>- Disclosure of conflicts of interest</li> </ul>	0-4 points



Applicability	<ul style="list-style-type: none"> <li>- Relevance to obesity detection</li> <li>- Discussion of clinical or practical implications</li> <li>- Assessment of implementation feasibility</li> </ul>	0-3 points
---------------	--	---------------

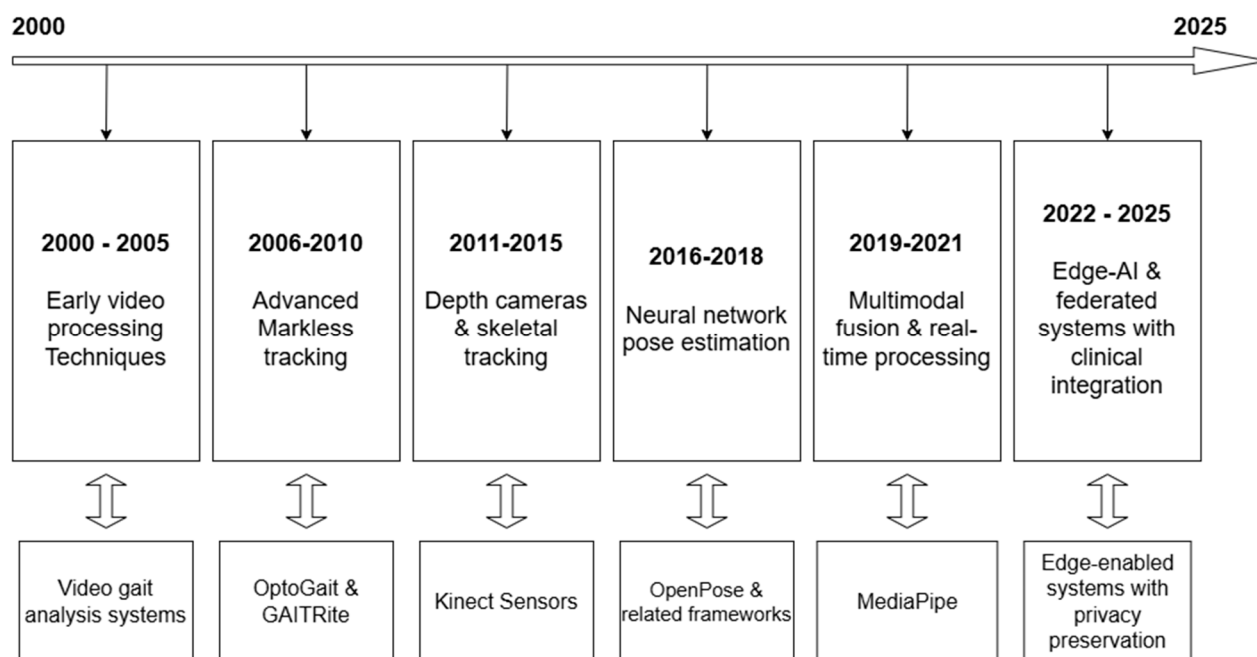
Total quality scores were categorized as follows:

- High quality: 20-24 points
- Moderate quality: 14-19 points
- Low quality: <14 points

We assessed each study, with discrepancies resolved through discussion. No studies were excluded based solely on quality assessment, but quality ratings were considered during data synthesis and interpretation.

### 2.5. Chronological Evolution Analysis

To capture the technological evolution in the field, we conducted a chronological analysis of optical sensing technologies for obesity detection from 2000 to 2025. Figure 1 illustrates this evolution.



**Figure 2.** Timeline of Key Developments in Optical Sensor Technologies for Obesity Detection.

### 2.6. Review Structure

The literature review was organized using a hybrid approach that combines technological taxonomies with application domains, enabling a comprehensive analysis of the field. This structure allows for both technical depth and application relevance:

#### 2.6.1. Primary Organization by Technology

##### Optical Gait Analysis Systems

- Light barrier technologies (e.g., OptoGait)
- Pressure-sensitive walkways (e.g., GAITRite)

- Video-based markerless systems
- Multi-camera setups

#### **Vision-Based Pose Estimation Frameworks**

- 2D pose estimation approaches (e.g., OpenPose, MediaPipe)
- 3D pose reconstruction methods
- Deep learning architectures (e.g., CNNs, transformers)
- Multi-person tracking systems

#### **Depth-Sensor Based Voxel Modeling**

- Structured light systems (e.g., first-generation Kinect)
- Time-of-flight sensors (e.g., Azure Kinect, RealSense)
- 3D body composition analysis
- Dynamic modeling approaches

#### **Hybrid and Multimodal Systems**

- Sensor fusion architectures
- Combined optical-inertial systems
- Multi-view integration approaches
- Ensemble methods

### 2.6.2. Secondary Organization by Application Focus

Within each technological category, studies were further organized according to their primary application focus:

#### **Biomechanical and Kinematic Analysis**

- Spatiotemporal gait parameters
- Joint angles and ranges of motion
- Center of mass trajectories
- Dynamic stability metrics

#### **Anthropometric Measurement and Validation**

- Body volume estimation
- Circumference measurements
- Body shape analysis
- Segmental proportions

#### **Obesity Classification and Risk Assessment**

- Algorithm development and validation
- Feature extraction methodologies
- Classification performance metrics
- Threshold determination

#### **Implementation and Deployment Frameworks**

- Clinical integration pathways
- Edge computing implementations
- Privacy-preserving architectures

- Real-world deployment considerations

This dual organizational structure enables the identification of both technological trends and application-specific challenges across the field of optical sensor-based obesity detection.

To summarize, this methodology ensures transparency and reproducibility of the literature review process. By adhering to PRISMA, we aim to strengthen the credibility of our findings and align with best practices required for publication in high-impact journals. Future research would benefit from the establishment of a shared repository of benchmark datasets for cross-validation and method comparison.

### 3. Optical Sensors Technologies for Gait Analysis in Obesity Detection

As obesity has become a global health concern, it is also associated with considerable motor impairments, particularly affecting gait and balance. Objective gait analysis techniques offer valuable tools for assessing these impairments and potentially identifying obesity-related gait alterations.

In this part, we examine the application of optical sensor-based systems, specifically image processing and floor sensor technologies, for gait analysis in individuals with obesity compared to normal-weight controls, drawing upon insights from the provided literature. We discuss the principles and hardware configurations of these systems, explore the gait biomarkers identified in the context of obesity, and analyze the technical advantages and limitations inherent in their application.

#### 3.1. Sensor Technologies Overview / Optical Gait Sensing for Obesity Detection

Gait analysis traditionally relies on either semi-subjective observations or objective measurements using various sensor technologies [7] [8]. Objective methods leverage technological advancements to quantify gait parameters with greater accuracy, exactitude, repeatability, and reproducibility compared to subjective assessments [7]. These objective techniques can be broadly categorized based on sensor placement: Non-Wearable Sensors (NWS) and Wearable Sensors (WS) [7]. Optical sensor-based systems primarily fall under the NWS category, requiring controlled laboratory settings where subjects walk along defined walkways equipped with sensors [7].

Optical sensor-based gait analysis relevant to obesity can be understood through the lens of floor sensor systems and image processing (video-based capture) technologies.

##### 3.1.1. Floor Sensor Systems

Floor sensor systems are a type of NWS where sensors are integrated into the floor or a walkway [7]. These systems measure gait by capturing data as the subject walks across them [7]. Two primary types are mentioned:

- Force Platforms: These systems utilize pressure or force sensors and moment transducers to measure the force vector applied during gait [7]. They can measure the 3D Ground Reaction Force (GRF) and moments involved in locomotion [8]. While highly accurate, they often require the subject to make contact with a specific area for correct measurement [7].
- Pressure Measurement Systems: Similar to force platforms, these systems quantify the center of pressure but do not directly measure the force vector [7]. They use arrays of sensitive cells (capacitive/resistive) to record plantar pressure distribution over time, revealing foot loading patterns and Center of Pressure (CoP) progression [8]. Examples include pressure sensor mats and platforms [7].

Floor sensor systems provide objective data on force-related parameters and gait phases [7]. They are considered gold standards for gait measurements, offering excellent quality data with high accuracy and repeatability, but are costly, bulky, and require a specialized workforce [8].

### 3.1.2. Optical Timing Systems

While not explicitly detailed as a separate category distinct from other floor or image-based systems in the sources, the description of systems like OptoGait suggests a form of optical timing and measurement [9].

OptoGait is described as a portable photoelectric cell system used for clinical assessment of static and dynamic foot pressures and quantifying spatio-temporal parameters. It works by measuring foot movements and space-temporal relationships using photoelectric cells [12]. The system is noted for its reliability in clinical assessment [9]. While technically leveraging photoelectric principles rather than image processing or traditional force/pressure plates, it functions similarly to some floor-based systems by assessing gait on a walkway and is often used to derive similar spatio-temporal parameters.

### 3.1.3. Video-Based Capture (Image Processing)

Image processing (IP) techniques utilize cameras to capture and analyze gait [7]. These systems extract essential gait features from images [7]. They range from single-camera systems to more complex multi-camera setups [7,8].

- **Marker-Based Systems:** These optical motion capture systems track targeted joints and orientations using reflective markers placed on the body [8]. They use multi-camera stereophotogrammetric video systems to compute the 3D localization of these markers, determining joint positions and body segment orientations [8].
- **Markerless Systems:** These systems use a human body model and image features to determine shape, pose, and joint orientations without the need for markers [8]. Recent work utilizes computer vision techniques and deep neural networks to extract 2D skeletons from images for gait analysis, even exploring privacy-preserving methods by processing encrypted images [10]. Examples include systems based on single cameras, Time of Flight sensors, Stereoscopic Vision, Structured Light, and IR Thermography [7].

Image processing systems allow for individual recognition and segment position analysis [7]. They offer advantages like relatively simple equipment setup for single cameras but can involve complex analysis algorithms and high computational costs for more advanced configurations [7]. They require controlled laboratory environments [8].

In summary, optical sensor-based systems for gait analysis in the context of obesity primarily involve floor-mounted force/pressure sensors and camera-based image processing systems. These NWS provide objective, quantitative data in a controlled setting, although they differ in the specific parameters they measure (forces vs. kinematics) and their technical complexity and cost [7,8]. The OptoGait system, a photoelectric cell-based system, also falls under this umbrella of fixed-location measurement systems used for gait assessment [9].

## 3.2. Applications in Obesity Context: Identified Biomarkers

Obesity is clearly linked to motor impairments, including deficits in gait and balance [11]. Individuals with obesity exhibit differences in movement and gait compared to those with normal weight, contributing to an increased risk of falls and stumbling [11]. Gait analysis using objective methods, including optical sensor-based systems, is applied to quantify these differences and identify specific gait alterations or biomarkers associated with obesity [5].

The goal is to capture gait and balance impairment in individuals with obese BMI and relate it to specific parameters [11]. While one source mentions findings that did not show significant differences in cadence, gait speed, stride duration, daily step count, or double support time between normal and obese BMI categories, it also notes that these findings diverge from existing literature [5].

Other sources and the research questions themselves highlight the expectation and investigation of such differences [5,11].

Typical gait parameters investigated in the context of obesity using objective systems include spatiotemporal parameters, kinematics, and kinetics [5,8,11].

### 3.2.1. Spatiotemporal Parameters

These include metrics like gait speed, step length, stride length, cadence, step width, step angle, step time, swing time, stance time, and double support time [5,7]. Studies aim to investigate variances in these parameters between obese and normal weight groups. Koinis et al. suggests that increasing BMI is associated with decreased gait speed and that obesity significantly increases the likelihood of falls [5]. Koinis et al. notes that people with obesity may experience up to a 15% reduction in gait speed and a 25% decrease in step length compared to those with normal BMI, although their own study did not find significant differences in some parameters[5]. Spatiotemporal parameters, especially walking speed and step length, are considered clinically important indicators [7].

### 3.2.2. Kinematics

This describes the movement of joints and body segments, including range of motion and segment acceleration [8,12]. While less explicitly detailed in relation to optical image systems and obesity in the provided excerpts compared to spatiotemporal parameters, biomechanical studies of obesity-related gait do investigate joint mechanics [5,7,11,13]. Image processing systems (marker-based and markerless) are capable of measuring joint angles and segment position/orientation [7,8].

### 3.2.3. Kinetics

This focuses on the forces and moments that cause movement, such as Ground Reaction Forces (GRF), muscle force, and joint momentum [8,12]. Floor sensor systems, particularly force platforms, are designed to measure GRF [7,8]. These kinetic parameters provide insight into the biomechanical effects of increased body mass on the musculoskeletal system during gait[5,7,13].

In addition to these quantitative parameters, gait analysis can also reveal qualitative aspects and patterns, such as gait symmetry and postural balance [7,8,13]. While specific findings on gait asymmetry directly measured by optical sensors in obese individuals aren't detailed across the sources, a study on overweight and obese children mentions assessing pelvic symmetry indices using a wearable system (BTS G-WALK, which uses inertial sensors) [13]. Postural balance is a key problem associated with conditions affecting gait, including obesity [7,8,11]. Gait and balance analysis are crucial for understanding locomotor and functional impairments [8].

Therefore, optical sensor-based systems, particularly floor sensors and camera-based systems, are used to objectively measure spatiotemporal, kinematic, and kinetic parameters that serve as biomarkers of obesity-related gait impairments, including potential changes in speed, step/stride length, timing, forces, joint movements, and overall gait pattern and stability [5,7,8,11].

## 3.3. Technical Advantages and Limitations

Objective gait analysis systems, including those based on optical sensors, offer significant advantages over traditional semi-subjective methods by providing accurate and quantitative data [7]. However, they also present technical limitations, particularly when considering their application in diverse settings and populations, such as individuals with obesity.

### 3.3.1. Precision vs. Portability Trade-Offs

Optical sensor-based NWS systems, operated in controlled laboratory environments, are considered gold standards for gait measurement due to their high accuracy and repeatability [8]. For instance, GRF plates offer high accuracy with minimal load error, and pressure sensor mats can achieve high recognition rates [1]. These systems provide precise data on a wide range of parameters

simultaneously [2]. However, their precision comes at the cost of portability. NWS systems are bulky, require specialized facilities, and are not suitable for monitoring gait during everyday activities outside the laboratory [7,8]. In contrast, wearable sensor (WS) systems offer portability and the ability to monitor gait in real-world settings over long periods, but often have reduced accuracy and reliability compared to NWS [7,8]. While WS are not the focus of this review section on optical sensors, this contrast highlights the inherent trade-off: high precision is typically found in non-portable, controlled NWS (like optical systems), while portability is the domain of WS with variable accuracy [8,14].

### 3.3.2. Environmental Dependencies, Calibration Needs, and Other Factors

Optical sensor-based systems can be sensitive to environmental factors and require careful setup and calibration.

- **Controlled Environment:** NWS, including image processing and floor sensors, require controlled research facilities. Subjects must walk on a clearly marked walkway [7].
- **Calibration:** Both image processing and floor sensor systems require calibration. For instance, stereoscopic vision systems involve complex calibration, and structured light systems also require calibration [7]. While the sources don't detail the specific calibration requirements for obese subjects, increased body size or altered gait patterns could potentially influence calibration procedures or accuracy.
- **Surface Sensitivity and Footwear Interference:** Floor sensor systems are directly affected by the interaction between the foot and the sensing surface [7,8]. The type of footwear worn can influence pressure distribution and force measurements, potentially acting as an interference or requiring standardized footwear [2].
- **Subject-Specific Variance:** While not unique to optical systems, individual variations in gait patterns are inherent. In the context of obesity, larger body mass significantly affects biomechanics and gait patterns [5,7,13]. Accurately capturing these subject-specific variations requires robust measurement techniques. Image processing systems that track body segments or skeletons may need to account for differences in body shape and soft tissue movement in obese individuals [10].

### 3.3.3. Limitations Specific to System Types:

#### **Floor Sensors**

Force platforms require the subject to contact the center of the plate for correct measurement. Pressure sensor mats and platforms have limitations of space, indoor measurement, and depend on the patient's ability to make contact with the platform.

#### **Image Processing**

Single camera systems have simple equipment but require complex analysis algorithms. Stereoscopic vision has complex calibration and high computational cost. Time of flight systems can have problems with reflective surfaces. IR Thermography requires considering emissivity, absorptivity, reflectivity, and transmissivity of materials. Extracting parameters like step length from image-based systems can sometimes be more accurate than methods used in some WS systems.

In summary, optical sensor-based gait analysis systems, as NWS technologies, offer high precision and the ability to capture detailed kinematic and kinetic data in a controlled environment [7,8]. However, they are non-portable, require specialized facilities and expertise, and can be sensitive to environmental factors, calibration specifics, and the interaction between the subject's foot/body and the sensing system [7,8,15]. Applying these systems to study gait in individuals with obesity

necessitates careful consideration of how increased body mass and potential gait alterations might influence data acquisition and interpretation [5,7,13].

**Table 4.** Comparative Analysis for Optical Sensor-Based Gait Analysis Systems (Based on Provided Sources).

Feature / System Type	Principles & Hardware Setup	Applications (Obesity Context)	Technical Advantages	Technical Limitations
Floor Sensors: Force Platforms	Measure 3D force vector, pressure, moment using sensors/transducers in floor.	Measure GRF, potentially revealing kinetic adaptations to increased body mass in obesity. Assess gait phases.	High accuracy (e.g., $\pm 0.1\%$ load error). Objective, quantitative data. Gold standard [8].	Requires controlled lab. Requires subject to contact center of plate. Bulky, costly, requires expertise [7,8,15]. Non-portable. Footwear can interfere.
Floor Sensors: Pressure Systems	Measure plantar pressure distribution and CoP using sensor arrays in floor mats/platforms.	Assess foot loading patterns and weight distribution during gait in obese individuals. Assess gait phases, step detection.	Measures plantar pressure patterns. Can have high recognition rates (80%) [7]. Easy setup in insoles (WS variant, but principle similar to NWS).	Limitations of space, indoor measurement. Patient must make contact with platform. Highly nonlinear response for insole type. Non-portable (mats/platforms). Surface/footwear sensitive.
Optical Timing (e.g., OptoGait)	Uses photoelectric cells along a walkway to measure foot movements and spatio-temporal timing.	Quantify spatio-temporal parameters (speed, timing, lengths) which are altered in obesity [5]. Reliable for clinical assessment [9].	Portable compared to larger NWS. Reliable for spatio-temporal measures [9].	Limited to spatio-temporal parameters [9]. Requires specific walkway setup. Can be sensitive to ambient light/interference (inferred from photoelectric principle).
Video-Based Capture: Marker-Based	Uses multi-camera stereophotogrammetry to track reflective markers on body segments.	Measure 3D kinematics (joint angles, segment position/orientation), revealing changes in movement patterns due to obesity's	High accuracy for kinematic measures [14]. Detailed 3D motion data.	Requires controlled lab with multiple cameras. Complex setup and calibration. Markers can be displaced by soft tissue or

		biomechanical effects [5,7,13]. Assess gait phases.		movement. Costly, requires expertise [8,15]. Non-portable.
Video-Based Capture: Markerless	Uses human body models and image features (e.g., 2D skeletons) from cameras (single, ToF, stereo, structured light, IRT).	Measure kinematics (segment position, joint angles), assess gait phases, potentially gait recognition or abnormal pattern detection. Useful for studying biomechanical changes in obesity.	Non-invasive. Can potentially work with less equipment (single camera). Progress in privacy preservation [10].	Accuracy can vary (moderate-poor for spatio-temporal in some WS applications, but NWS generally better [14]). Complex analysis algorithms (single camera). Complex calibration/high computational cost (stereo vision). Issues with reflective surfaces (ToF). Requires specific environmental conditions (IRT). Non-portable (NWS setups). [16]

In this section, we highlighted the critical role of objective gait analysis in understanding and quantifying the motor impairments associated with obesity. Optical sensor-based systems, including floor sensor technologies (force platforms, pressure systems) and video-based capture (marker-based and markerless image processing), represent key non-wearable sensor (NWS) approaches used in this field [7,8]. These systems are capable of measuring important gait biomarkers such as spatiotemporal parameters, kinematics, and kinetics, which are known to be altered by increased body mass and contribute to mobility issues and fall risk in individuals with obesity [5,7].

While NWS optical systems offer high accuracy and the ability to collect comprehensive data in controlled settings, they are limited by their lack of portability, high cost, need for specialized expertise, and susceptibility to environmental and subject-specific factors [7,8,15]. Despite these limitations, they remain valuable tools for detailed clinical and research assessments of gait mechanics in obesity.

Future advancements, particularly in areas like miniaturization, power efficiency, and sophisticated algorithms, are focused on improving wearable technologies to potentially bridge the



gap in measurement capacity and accuracy with NWS, enabling long-term, real-world gait monitoring. However, for detailed, high-precision laboratory-based analysis, optical sensor systems continue to play a significant role in uncovering the complex interplay between obesity and gait dynamics. Further research utilizing these objective techniques is essential for refining our understanding of obesity-related gait abnormalities and developing targeted interventions.

## 4. Markerless Video-Based Pose Estimation Technologies

Markerless pose estimation represents a revolutionary approach to human motion analysis, enabling the extraction of kinematic data without the need for physical markers attached to subjects. This technology has seen rapid advancement in recent years, primarily driven by developments in computer vision and deep learning. Unlike traditional marker-based motion capture systems that require specialized hardware and controlled laboratory environments, markerless systems operate with standard cameras in diverse settings, making them accessible for widespread applications in healthcare, sports science, biomechanics research, and human-computer interaction. This part of the review examines current markerless video-based pose estimation technologies, focusing on algorithms, validation, challenges related to body morphology diversity, and advancements in hybrid sensing approaches.

### 4.1. Key Algorithms and Platforms

#### 4.1.1. OpenPose

OpenPose represents one of the pioneering deep learning-based frameworks for real-time multi-person human pose detection. Developed at Carnegie Mellon University, it revolutionized the field by enabling the simultaneous detection of multiple individuals within a single image or video frame. OpenPose employs a bottom-up approach that first detects body parts across the entire image and then associates them to form complete human skeletons.

The architecture of OpenPose is built upon a multi-stage convolutional neural network (CNN) that processes images through two main branches: one for body part detection and another for part association. This two-branch approach enables the system to maintain high accuracy even when multiple people appear in the scene with overlapping body parts. The network generates confidence maps for each body part location and part affinity fields (PAFs) that encode the degree of association between parts, allowing the system to determine which body parts belong to the same person.

OpenPose can jointly detect human body, foot, hand, and facial keypoints, providing a comprehensive representation of human pose. The standard model identifies 25 body keypoints, including major joints like shoulders, elbows, wrists, hips, knees, and ankles, as well as facial landmarks. Extended models incorporate additional keypoints for hands and detailed facial features, resulting in a total of 135 keypoints per person when using the full model.

The versatility of OpenPose has led to its application across diverse domains. In biomechanics research, it has enabled the analysis of sports performance without interfering with athletes' natural movements. In 2020, Nakano et al. developed a 3D markerless motion capture technique using OpenPose with multiple synchronized cameras to evaluate motor performance tasks including walking, jumping, and ball throwing [17]. They found that approximately 47% of measurements had mean absolute errors below 20mm compared to marker-based systems, with 80% below 30mm [17].

#### 4.1.2. MediaPipe

MediaPipe Pose is another significant deep learning-based framework for human pose estimation, developed by Google. Unlike OpenPose's bottom-up approach, MediaPipe typically employs a top-down methodology that first detects persons in the image and then estimates the pose for each detected individual. This approach generally works well when the number of people in the

scene is limited, making it particularly suitable for applications focusing on a single subject or a few individuals.

MediaPipe Pose Estimation is based on the BlazePose architecture, which was specifically designed for real-time performance on mobile devices [18]. The system provides 33 3D keypoints in real-time, representing a superset of the 17 keypoints from the COCO dataset (commonly used in many pose estimation systems). These additional points provide more detailed tracking of the face, hands, and feet, enhancing the granularity of pose information [18]. The pipeline of MediaPipe Pose first detects a person in the image using a face detector and then predicts the keypoints, assuming that the face is always visible [18].

A distinctive feature of MediaPipe is its optimized performance for mobile deployment. On devices like the Samsung Galaxy S23 Ultra with the Snapdragon 8 Gen 2 chipset, the inference time can be as low as 0.826 ms, with a peak memory range of 0-1 MB [19]. This exceptional efficiency makes MediaPipe an excellent choice for real-time applications on edge devices where computational resources are limited.

MediaPipe Pose is primarily designed for fitness applications involving a single person or a few people in the scene [18]. Its applications include yoga pose correction, fitness tracking, physical therapy, and gesture-based interfaces. The framework is easily accessible through Python packages and can be configured to run on cloud-hosted devices using platforms like the Qualcomm AI Hub [19].

#### 4.1.3. DeepLabCut

DeepLabCut represents a different approach to pose estimation, originally developed for markerless tracking of animals in research settings. Created by Mathis et al., DeepLabCut leverages transfer learning to achieve high-performance pose estimation with relatively small training datasets, making it particularly valuable for specialized applications where large annotated datasets may not be available [20].

The architecture of DeepLabCut was initially inspired by DeeperCut, a state-of-the-art algorithm for human pose estimation by Insafutdinov et al. [21], which inspired the name for the toolbox. However, since its inception, DeepLabCut has evolved substantially, incorporating various backbone networks including ResNets, MobileNetV2, EfficientNets, and the custom DLCRNet backbones. This flexibility in network architecture allows users to balance accuracy and computational efficiency based on their specific requirements.

A key strength of DeepLabCut is its ability to achieve high accuracy with limited training data, typically requiring only a few hundred labeled frames to generate reliable pose estimates for novel videos [20]. This is achieved through transfer learning, where pre-trained networks (typically trained on ImageNet) are fine-tuned for specific pose estimation tasks. The developers have demonstrated that this approach works effectively across species including mice, flies, humans, fish, and horses.

In addition to its 2D pose estimation capabilities, DeepLabCut also supports 3D pose reconstruction using multiple cameras or even from a single camera with appropriate training data [22]. The framework has been extended to support real-time processing through DLClive, enabling applications that require immediate feedback based on pose information [22].

While DeepLabCut was originally developed for animal tracking, its principles and approaches have been successfully applied to human subjects as well [23]. The framework is particularly valuable in research contexts where custom keypoint definitions may be needed, or where the specifics of the application differ from the standard human pose estimation use cases [24].

### 4.2. Validation and Accuracy

#### 4.2.1. Comparison with Gold Standard Systems

The validation of markerless pose estimation systems against gold standard marker-based motion capture is essential for establishing their reliability in scientific and clinical applications.

Optical marker-based systems, such as Vicon or OptiTrack, remain the reference standard in biomechanics research due to their sub-millimeter accuracy in controlled environments.

A comprehensive evaluation of OpenPose-based markerless motion capture was conducted by Nakano et al., comparing it with optical marker-based systems during various motor tasks including walking, countermovement jumping, and ball throwing [17]. The study employed multiple synchronized cameras to reconstruct 3D poses from OpenPose's 2D estimates and compared the resulting joint positions with those measured by a marker-based system. The differences were quantified using mean absolute error (MAE) between corresponding joint positions [17].

The results revealed that approximately 47% of all calculated mean absolute errors were below 20 mm, and 80% were below 30 mm, indicating reasonable accuracy for many applications [17]. However, approximately 10% of errors exceeded 40 mm, primarily due to failures in OpenPose's 2D tracking, such as incorrectly recognizing objects as body segments or confusing one body segment with another [17]. These findings suggest that while markerless systems can approach the accuracy of marker-based systems for many applications, they still face challenges in robustly tracking all body segments across diverse movements and viewing conditions [25]

The accuracy of markerless systems varies considerably across different joints and movement types. Generally, larger and more visible joints such as the shoulders, hips, and knees tend to be tracked more reliably than smaller joints like the wrists, ankles, and fingers. Additionally, movements that involve rapid motion, occlusion, or unusual poses can challenge the performance of current algorithms, leading to increased error rates [26].

It's important to note that while absolute position accuracy may still lag behind marker-based systems, many applications primarily require accurate relative motion patterns or joint angles, which markerless systems can often provide with sufficient reliability. This makes them viable alternatives for applications where the convenience of markerless tracking outweighs the need for the highest possible accuracy.

#### 4.2.2. Comparison with IMU Systems

Inertial Measurement Units (IMUs) represent another approach to motion capture that relies on wearable sensors containing accelerometers, gyroscopes, and sometimes magnetometers to track segment orientations. While not entirely "markerless" in the strictest sense (as they require sensors to be attached to the body), IMUs offer an alternative to optical systems that have gained popularity for field-based measurements.

Recent research has compared both IMU and markerless video-based methods against optoelectronic systems. A pilot validation study examined these approaches during simulated surgery tasks, where traditional marker-based systems would be impractical [27]. The findings indicated that the IMU method demonstrated root mean square errors (RMSE) of 2.1 to 7.5 degrees with intraclass correlation coefficients (ICC) ranging from 0.53 to 0.99, while the markerless method showed higher errors of 5.5 to 8.7 degrees RMSE with ICCs between 0.31 and 0.70 [27]. Based on these results, the researchers recommended the IMU method over the markerless approach for the specific context of measuring neck and trunk movements during surgery [27].

However, it's important to consider the relative strengths and weaknesses of each approach. IMUs excel at capturing segment orientations and can operate without line-of-sight constraints, making them suitable for complex environments with occlusions. However, they struggle with position drift over time and require careful calibration. Markerless video systems, in contrast, can provide absolute position information without drift but require continuous visibility of body segments to the cameras.

The complementary nature of these technologies has led to increasing interest in hybrid systems that combine IMUs with video-based tracking to leverage the strengths of each approach. Such systems can use visual information to correct IMU drift while using IMU data to fill gaps during visual occlusions.

### 4.2.3. Body Morphology Effects on Detection

The impact of body morphology, particularly body mass index (BMI), on the accuracy of markerless pose estimation represents a significant challenge for these technologies. High BMI can affect pose estimation accuracy through several mechanisms. First, increased adipose tissue can change the visual appearance of joints, making their precise localization more difficult. Second, in individuals with higher BMI, certain joints may be partially occluded by soft tissue, reducing their visibility to the camera. Third, the standard body proportions assumed by many pose estimation algorithms may not accurately represent individuals with higher BMI, potentially leading to systematic errors in keypoint placement.

Limited research has directly quantified these effects, but clinical experience and preliminary studies suggest that pose estimation accuracy generally decreases as BMI increases, particularly for joints of the lower extremities. This creates a significant challenge for applications in healthcare settings, where individuals with higher BMI may be precisely those who would benefit most from motion analysis for conditions like osteoarthritis, diabetic gait disorders, or rehabilitation monitoring.

To address these challenges, several approaches have been proposed. One approach involves creating more diverse training datasets that include individuals across the full spectrum of body sizes and shapes. Another approach uses adaptive algorithms that can adjust their keypoint detection strategies based on the detected body morphology. Some researchers have also explored the use of additional sensors, such as depth cameras, to provide supplementary information that can improve joint localization in challenging cases.

The ability to accurately track movements across diverse body morphologies remains an important frontier for markerless pose estimation research, with implications for the equity and inclusivity of these technologies in healthcare and other domains.

## 4.3. Obesity-Related Gait Signatures

### 4.3.1. Technical Challenges

Markerless pose estimation faces several technical challenges when applied to individuals with obesity, particularly in the context of gait analysis. The first major challenge is joint occlusion, which occurs when adipose tissue or limb positioning prevents clear visual access to joint centers. This is especially problematic for the hip joints, which may be obscured by abdominal or thigh tissue, and for the knees, which can be partially hidden during certain phases of the gait cycle.

Over-segmentation represents another challenge, where the algorithm incorrectly identifies multiple keypoints where only one should exist. This can occur when the visual appearance of body segments in high-BMI individuals differs significantly from the training data used to develop the pose estimation model. For example, the algorithm might mistakenly identify multiple knee joints due to the different contour of the leg in individuals with higher BMI.

Signal processing adaptations have been developed to address these challenges. These include temporal filtering approaches that maintain continuity of joint trajectories based on biomechanical constraints, preventing physically impossible jumps in joint positions between frames. Some systems also incorporate anatomical constraints and body-specific calibration procedures to adapt their models to individual body morphologies.

Multi-view approaches can significantly mitigate occlusion issues by providing alternative angles from which to observe partially hidden joints. When a joint is occluded from one camera's perspective, it may be visible from another, allowing the system to maintain tracking. Advanced systems can dynamically weight the confidence of detections from different cameras based on their viewing angle relative to each body segment.

Addressing these technical challenges is essential for developing inclusive motion analysis technologies that can serve diverse populations. The most promising approaches combine algorithmic improvements with hardware solutions like strategic camera placement to maximize visibility of key anatomical landmarks.

#### 4.3.2. Biomechanical Alterations

Obesity is associated with several characteristic alterations in gait biomechanics that pose estimation systems must accurately capture to provide clinically relevant information. Understanding these patterns is essential both for developing more robust tracking algorithms and for interpreting the resulting kinematic data in clinical contexts.

Altered joint angle trajectories represent one of the most significant gait modifications in individuals with obesity. Typically, these include reduced knee flexion during swing phase, decreased hip extension during late stance, and modified ankle kinematics throughout the gait cycle. These alterations are believed to result from a combination of increased joint loading, altered muscle function, and adaptations to maintain stability with changed body mass distribution.

Increased trunk lean is another common characteristic of gait in individuals with higher BMI. This forward inclination of the trunk shifts the center of mass anteriorly, potentially reducing the muscular effort required to initiate forward progression during walking. Accurately quantifying trunk lean is important for assessing energy expenditure during gait and for understanding compensatory mechanisms that may increase risk for back pain or other musculoskeletal issues.

Lateral sway patterns also differ in individuals with obesity, with typically increased mediolateral center of mass displacement during walking. This increased lateral movement requires additional stabilizing mechanisms and may contribute to higher energy costs of walking. Capturing these subtle movements requires pose estimation systems with high accuracy in tracking the relative positions of the pelvis, lower extremities, and trunk.

Markerless systems must be capable of accurately measuring these biomechanical alterations to provide clinically meaningful assessments. Validation studies specifically examining the accuracy of these systems in capturing obesity-related gait signatures are limited but represent an important area for future research.

#### 4.3.3. Clinical Applications

Despite the challenges, markerless pose estimation offers significant potential for clinical applications related to obesity and associated movement disorders. The non-invasive nature of these systems makes them particularly valuable for longitudinal monitoring, where repeated assessments are needed to track changes over time.

In weight management programs, objective quantification of gait parameters can provide valuable feedback on the functional improvements resulting from weight loss. Parameters such as step length, walking speed, joint ranges of motion, and stability measures can demonstrate functional gains that may motivate continued adherence to intervention programs. Markerless systems enable these measurements to be taken in clinical settings without the time-consuming application of markers or specialized equipment.

For surgical interventions such as bariatric surgery or joint replacements, markerless motion analysis can help document functional outcomes and guide rehabilitation strategies. The ability to conduct these assessments quickly and easily facilitates their integration into routine clinical care, rather than being limited to specialized research settings.

Telehealth applications represent another promising domain, where markerless systems using standard webcams could enable remote assessment of movement function. This could be particularly valuable for monitoring patients in rural or underserved areas where access to specialized gait laboratories is limited.

As these technologies continue to improve in accuracy and robustness across diverse body morphologies, their integration into standard clinical care pathways for obesity and related conditions becomes increasingly feasible, potentially transforming the assessment and management of movement-related complications.

#### 4.4. Depth and Hybrid Systems

#### 4.4.1. RGB-D Framework

RGB-D systems combine traditional color images (RGB) with depth information (D), creating a more comprehensive representation of the 3D scene. While standard RGB cameras capture only the visual appearance of subjects, depth sensors provide direct measurements of the distance between the sensor and each point in the scene. This additional dimension of information can significantly enhance the accuracy and robustness of pose estimation, particularly in challenging scenarios involving occlusions or unusual body positions.

The Microsoft Kinect V2 represents one of the most widely used RGB-D platforms for human motion capture. It combines a standard RGB camera with an infrared time-of-flight depth sensor that provides pixel-wise distance measurements. The integration of depth data allows the system to disambiguate between overlapping body parts and more accurately localize joints in 3D space, even when their appearance in the RGB image alone might be ambiguous.

The processing pipeline for RGB-D pose estimation typically involves several stages. First, the depth information is used to segment the human figure from the background. Next, the segmented depth map is processed to identify body parts using techniques such as random decision forests or deep learning. Finally, a skeletal model is fitted to these detected body parts, taking into account both the RGB appearance and the 3D structure provided by the depth data.

More recent approaches have incorporated deep learning methods that can jointly process RGB and depth information. These networks are trained to leverage complementary cues from both modalities: appearance features from RGB images and structural information from depth maps. This fusion of information sources has proven particularly effective for robust pose estimation in complex real-world environments.

#### 4.4.2. Accuracy Improvements

The incorporation of depth information provides several significant accuracy improvements for pose estimation, especially in challenging scenarios. First, depth data helps resolve ambiguities in the RGB image by providing direct 3D information about the spatial arrangement of body parts. This is particularly valuable when body parts overlap from the camera's perspective, which can confuse RGB-only systems.

Second, depth sensors are generally less sensitive to lighting variations than RGB cameras, making them more robust for applications in environments with inconsistent or poor lighting. While strong infrared interference can affect depth sensors, they generally provide more stable measurements across varying ambient light conditions than color-based approaches alone.

Third, depth information facilitates more accurate background segmentation, helping to isolate the human figure from complex environments. This is especially valuable in cluttered scenes where color-based segmentation might struggle to distinguish between the subject and visually similar background elements.

Quantitative studies have demonstrated these advantages, with RGB-D systems typically showing reduced average joint position errors compared to RGB-only approaches when evaluated against marker-based ground truth. The magnitude of improvement varies by joint, with the greatest benefits often seen for joints that are frequently occluded or that lack distinctive color features.

However, it's important to note that depth sensors have their own limitations, including more restricted range, higher power consumption, and typically lower resolution than RGB cameras. These considerations are particularly relevant for mobile or wearable applications where power and computational resources may be constrained.

#### 4.4.3. Real-World Applications

RGB-D systems have found applications across numerous domains where robust pose estimation in uncontrolled environments is required. In clinical settings, these systems enable functional movement assessment without the need for markers, facilitating the integration of motion

analysis into routine care. Applications include gait assessment, balance evaluation, and rehabilitation monitoring, where the system can provide immediate feedback on movement quality and progress.

Home monitoring represents another growing application area, where RGB-D sensors can track movements over extended periods in naturalistic environments. This enables longer-term assessment of mobility patterns and functional status, which may be more representative of real-world capabilities than brief assessments in clinical settings. Privacy concerns in home monitoring can be mitigated by processing data locally and extracting only anonymous skeletal data rather than storing raw RGB images.

Public space analysis for ergonomics, safety, and accessibility represents a third application domain. Here, RGB-D systems can analyze how diverse individuals interact with built environments without requiring individual consent for marker placement. This supports the development of more inclusive design standards that accommodate the full range of human body sizes and movement capabilities.

The continued miniaturization and cost reduction of depth sensing technologies promises to further expand these applications. Emerging systems incorporate depth sensing directly into mobile devices or wearable cameras, enabling pose estimation in increasingly diverse and dynamic environments while maintaining user privacy through on-device processing of sensitive data.

Markerless video-based pose estimation technologies have advanced rapidly in recent years, driven by breakthroughs in deep learning and computer vision. Systems like OpenPose, MediaPipe, and DeepLabCut provide accessible frameworks for human motion analysis across diverse applications, from clinical assessment to sports performance and human-computer interaction.

Validation studies against gold standard marker-based systems indicate that markerless approaches can achieve reasonable accuracy for many applications, with the majority of joint position errors falling below 30mm in controlled conditions. However, challenges remain in tracking rapid movements, handling occlusions, and accurately capturing the movements of individuals whose body morphologies differ significantly from those represented in training datasets.

The impact of body morphology, particularly higher BMI, on pose estimation accuracy remains an important consideration for clinical applications. Technical challenges including joint occlusion and over-segmentation can affect the reliable tracking of obesity-related gait signatures such as altered joint trajectories, increased trunk lean, and modified lateral sway patterns. Addressing these challenges requires both algorithmic improvements and hardware solutions.

The integration of depth sensing with RGB cameras in hybrid systems offers promising improvements in robustness and accuracy, particularly in complex real-world environments. These RGB-D systems provide complementary information that enhances joint localization, improves robustness to lighting variations, and facilitates better segmentation of human figures from cluttered backgrounds.

Looking forward, the continued development of markerless pose estimation technologies promises to democratize access to human movement analysis, enabling applications that were previously confined to specialized laboratories to be deployed in clinical settings, homes, and public spaces. This expanded access has the potential to transform our understanding of human movement across diverse populations and environments, ultimately contributing to improved healthcare, enhanced performance, and more inclusive design of physical spaces and interfaces.

## 5. 3D Human Voxel Modeling and Anthropometric Estimation

The increasing availability of consumer-grade depth sensors has sparked significant research interest in 3D human body modeling and measurement extraction. This field intersects computer vision, machine learning, and anthropometry to develop methods for accurate body shape reconstruction and measurement estimation. This part of the review examines the current state of research in voxel-based human body modeling with a focus on anthropometric applications.

## 5.1. Sensor-Based 3D Body Reconstruction

### 5.1.1. Depth Sensing Technologies

The evolution of consumer-grade depth cameras has revolutionized 3D human body reconstruction techniques. Time-of-Flight (ToF) cameras like the Microsoft Kinect V2 measure depth by calculating the time taken for infrared light pulses to travel to an object and back. In contrast, stereoscopic cameras such as the Intel RealSense D435 derive depth information from parallax between two camera viewpoints. A comparative study by Chuang-Yuan et al. revealed that ToF cameras generally provide more accurate 3D point clouds for human body modeling than stereoscopic sensors, with the Kinect V2 outperforming the RealSense D435 in KinectFusion reconstruction quality [28].

Microsoft Kinect has been widely adopted for 3D body scanning applications due to its accessibility and reasonable accuracy. Weiss et al. demonstrated that the Kinect can produce 3D body models with measurement accuracy competitive with commercial body scanning systems costing orders of magnitude more [29]. Their approach combined low-resolution image silhouettes with coarse range data to estimate a parametric model of the body, achieving accurate results by combining multiple monocular views of a person moving in front of the sensor.

### 5.1.2. Voxel-Based Representation and Processing

Voxel-based representations serve as a fundamental approach to 3D body reconstruction from depth data. Voxels (volumetric pixels) discretize 3D space into regular grid cells, each containing occupancy information. Li et al. introduced a coarse-to-fine reconstruction method combining voxel super-resolution (VSR) with learned implicit representation [30]. Their approach first estimates coarse 3D models using a Pixel-aligned Implicit Function based on Multi-scale Features (MF-PIFu) extracted from multi-view images. The coarse model is then refined by implementing VSR through a multi-stage 3D convolutional neural network, which significantly enhances surface quality and geometric accuracy.

The resolution of voxel representations presents a critical trade-off between detail and computational efficiency. Research by Chuang-Yuan et al. demonstrated that increasing Kinect Fusion resolution from 128 to 512 voxels per meter yielded diminishing returns beyond 256 voxels per meter when scanning human subjects [28]. This finding suggests an optimal voxel resolution for body scanning applications that balances detail capture and processing requirements.

### 5.1.3. Single-View Versus Multi-View Reconstruction

While multi-view approaches traditionally yield superior reconstruction quality, recent advances have improved single-view reconstruction methods. Single-view reconstruction is particularly relevant for practical applications where multiple cameras or viewpoints are unavailable. The Pixel2Pose approach demonstrated by researchers employs neural networks trained on high-resolution depth and intensity images from Microsoft Kinect to recover poses of multiple people in three dimensions from single-view time-of-flight data [31]. Despite the apparent low spatial resolution of single-view captures, their system transforms the sensor's rich ToF data into accurate 3D pose information after supervised training.

For enhanced accuracy, multi-view approaches remain superior. As demonstrated by Li et al., combining multiple viewpoints allows for the creation of complete 3D models with fewer occlusions and better surface quality [30]. Their method uses implicit representations, which enable memory-efficient training and produce high-resolution continuous decision boundaries, addressing the challenge of limited resolution in voxel grids.

### 5.1.4. Parametric Body Models



Statistical parametric body models have emerged as powerful tools for robust 3D reconstruction from noisy and incomplete sensor data. The SCAPE (Shape Completion and Animation for PEople) model, as utilized by Weiss et al., factors 3D body shape variations from pose variations [29]. This approach enables the estimation of a single consistent body shape while allowing pose to vary, making it ideal for reconstructing consistent 3D human models from multiple partial views.

Parametric models provide a structured way to represent human body variation across a population and facilitate regression from sparse measurements to complete body shape. This capability is particularly valuable for anthropometric applications, as demonstrated by Tsoli et al., who used these models to predict body measurements with greater accuracy than state-of-the-art methods [32].

## 5.2. Applications in Body Composition Analysis

### 5.2.1. Anthropometric Measurement Extraction

3D voxel representations of the human body enable automated extraction of anthropometric measurements traditionally performed by human anthropometrists. Tsoli et al. developed a model-based anthropometry approach that fits a deformable 3D body model to scan data and then extracts features including limb lengths, circumferences, and statistical features of global shape [32]. Their method demonstrated superior accuracy compared to traditional landmark-based approaches, particularly when integrating information from multiple scans of a person in different poses.

The extraction of perimeter values from 3D scans has proven particularly valuable for body composition assessment. Alexa's research at Philips showed that models based on perimeter values around thigh, waist, arm, and neck could accurately predict body fat percentage, achieving an RMSE of 2.22% when compared to reference methods in pregnant women [33]. This approach demonstrates how geometric features from 3D body scans can translate into clinically relevant body composition metrics.

### 5.2.2. Waist-to-Hip Ratio and Volumetric Indices

The waist-to-hip ratio (WHR) represents a powerful predictor of health risks associated with fat distribution. LeanScreen technology has demonstrated the capability to calculate WHR quickly using 2D photographic methods combined with 3D body modeling techniques [34]. This approach exemplifies how even partial 3D reconstruction can yield clinically relevant anthropometric indices.

Volumetric metrics derived from voxel-based 3D body models provide comprehensive assessments of body composition. Alexa's research demonstrated that even from single Kinect depth maps (thus an incomplete 3D representation), predictive models for body fat percentage could be developed [33]. Using a Lasso regression approach with features extracted from point cloud data, the study achieved a predictive model with adjusted  $R^2 = 0.72$  and RMSE = 8.02%. This finding illustrates how volumetric data, even when incomplete, can yield valuable body composition information.

### 5.2.3. Shape Descriptors and Curvature Analysis

Beyond simple circumference measurements, advanced shape descriptors derived from 3D voxel models provide nuanced insights into body composition. Laws et al. highlighted the value of curvature analysis for body composition assessment, establishing relationships between surface geometry characteristics and underlying tissue distribution [35]. These shape-based features complement traditional anthropometric measures and enhance the predictive power of body composition models.

The combination of local and global shape features yields more robust body composition assessments than either approach alone. Tsoli et al. demonstrated that combining localized measurements (such as limb lengths and circumferences) with statistical features describing overall body shape variation significantly improved the accuracy of measurement predictions [32]. Their

inclusion of both feature types allowed their system to account for individual variations in body shape that might not be captured by standard anthropometric measurements alone.

#### 5.2.4. Comparison with Traditional Methods

The accuracy of voxel-based body composition assessment relative to traditional methods represents a critical consideration for clinical adoption. Research comparing 3D scan-derived measures against Dual-energy X-ray absorptiometry (DXA), hydrostatic weighing, and Bioelectrical Impedance Analysis (BIA) has shown promising results. Astorino et al. were among the first to investigate body composition measured with DXA compared to 3D scans, finding significant correlations between the methods though not equivalent performance [33].

Alexa's study comparing Kinect-based fat percentage prediction against BIA measurements demonstrated the potential of depth sensing for body composition analysis while highlighting the need for further refinement [33]. The achieved RMSE of approximately 8% represents a substantial improvement over earlier attempts but still falls short of the accuracy provided by laboratory methods, suggesting that 3D scanning technologies may serve as convenient screening tools rather than diagnostic replacements in their current form.

### 5.3. Gait Integration Possibilities

#### 5.3.1. Morphology-Locomotion Relationships

The integration of 3D body modeling with gait analysis offers powerful insights into the relationship between body morphology and movement patterns. The ability to accurately capture both static body shape and dynamic locomotion provides a comprehensive framework for analyzing how morphological variations influence movement efficiency and pathology.

The Pixels2Pose system demonstrates how time-of-flight imaging can be leveraged for 3D pose estimation, creating a bridge between static body modeling and dynamic movement analysis [31]. By transforming sensor data into accurate 3D skeletal poses, such systems enable simultaneous assessment of body morphology and movement patterns, facilitating the study of their interrelationship.

#### 5.3.2. Biomechanical Analysis and Clinical Applications

Voxel-based body models provide precise segmentation of body parts, enabling detailed biomechanical analysis of gait. The accurate calculation of segment masses, centers of mass, and moments of inertia from volumetric data enhances the precision of inverse dynamics calculations and other biomechanical analyses. This integration has a particular relevance for clinical populations where body morphology deviations may contribute to movement abnormalities.

The parametric body models discussed by Black et al. and Tsoli et al. provide structured representations that can be animated according to captured motion data [29,32]. This capability enables researchers to simulate how body shape variations might influence movement mechanics, offering a powerful tool for both clinical assessment and rehabilitation planning.

#### 5.3.3. Longitudinal Monitoring and Intervention Assessment

The combination of 3D body modeling and gait analysis creates opportunities for comprehensive longitudinal monitoring of patients undergoing rehabilitation or weight management interventions. By capturing both morphological changes and associated alterations in movement patterns, clinicians can assess intervention efficacy more holistically than with either approach alone.

The relatively low cost and portability of consumer-grade depth sensors make this integrated approach accessible for clinical practice. The work by Weiss et al. demonstrating accurate body scanning with inexpensive commodity sensors highlights the potential for widespread adoption of

these technologies in clinical settings [29], potentially revolutionizing how clinicians monitor the interplay between morphological changes and movement adaptations during recovery or disease progression.

#### 5.4. Limitations

##### 5.4.1. Segmentation Errors and Depth Artifacts

Despite significant advancements, 3D body reconstruction from depth sensors remains susceptible to segmentation errors and depth artifacts. The KinectFusion technique, while powerful, can produce artifacts at the boundaries between the subject and background, particularly in environments with complex geometry or variable lighting conditions [28]. These artifacts can distort body shape reconstruction and compromise the accuracy of derived measurements.

Depth artifacts present challenges when scanning subjects with obesity. Alexa's research noted difficulties in accurately capturing body contours in participants with higher body fat percentages, as the increased tissue depth and surface curvature created shadows and occlusions in the depth map [33]. These artifacts can lead to systematic underestimation of body volume in subjects with obesity, potentially limiting the clinical utility of the technology for this population.

##### 5.4.2. Resolution and Surface Quality Limitations

The resolution of consumer-grade depth sensors imposes fundamental limitations on the detail level achievable in 3D body reconstruction. Li et al. addressed this challenge through their voxel super-resolution approach, which enhances the detail level of coarse 3D models [30]. However, this approach, while improving results, cannot fully compensate for the inherent resolution limitations of the source data.

Surface quality represents another significant limitation of current approaches. The point clouds generated by depth sensors often contain noise and irregularities that impact the smoothness and accuracy of the reconstructed surface. This limitation affects the precision of curvature-based shape descriptors and may reduce the accuracy of derived anthropometric measurements, particularly for small or subtle anatomical features.

##### 5.4.3. Posture Variability and Subject Positioning

Posture variability introduces significant challenges for 3D body reconstruction and measurement extraction. Even minor variations in standing posture can alter key anthropometric measurements and confound longitudinal comparisons. The approach by Weiss et al. using the SCAPE body model partially addresses this issue by allowing pose to vary while maintaining a consistent body shape [29], but the challenge persists in practical applications where standardized positioning may be difficult to achieve.

Subject positioning relative to the sensor also influences reconstruction quality. Alexa's research utilized five different predetermined angles between the subject and the Kinect plane to capture comprehensive body data [33]. This approach highlights the dependency of reconstruction quality on careful subject positioning, which may limit the practical applicability of the technology in uncontrolled environments.

##### 5.4.4. Clothing and Surface Appearance Effects

Clothing presents a significant confusing factor for 3D body reconstruction and anthropometric measurement. Loose or bulky clothing obscures true body contours, while tight clothing may compress soft tissues and alter measurements. Most research protocols, including those reviewed here, require standardized tight-fitting clothing to minimize these effects, limiting the ecological validity of the resulting measurements.

Surface appearance characteristics, including skin tone, reflectance properties, and texture patterns, can also affect depth sensing accuracy. Particularly for stereo-based systems like the RealSense D435, surfaces lacking texture detail may result in poor depth estimation [28]. While time-of-flight systems like the Kinect V2 are less affected by surface appearance, they remain sensitive to highly reflective or absorbent materials, potentially limiting their effectiveness across diverse populations.

#### 5.4.5. Accuracy Compared to Gold Standards

Despite promising results, voxel-based anthropometric measurements have not yet achieved the accuracy of gold standard methods like DXA or hydrostatic weighing. Alexa's research achieved an RMSE of approximately 8% for body fat percentage prediction from Kinect scans compared to BIA measurements [33], which themselves have limitations compared to gold standard methods. This accuracy gap suggests that current 3D scanning approaches may be better suited for tracking relative changes than for absolute diagnostic measurements.

The comparison conducted by Chuang-Yuan et al. between different depth cameras revealed that even under optimal conditions, depth sensing technologies introduce measurement errors that propagate through the reconstruction pipeline [28]. These fundamental limitations of sensing technology establish a ceiling on the achievable accuracy of derived anthropometric measurements, highlighting the need for continued technological advancement.

To summarize, Voxel-based 3D human body modeling from consumer-grade depth sensors represents a rapidly evolving approach to anthropometric measurement and body composition analysis. The techniques reviewed in this section demonstrate significant progress in addressing key challenges, including accurate shape reconstruction from incomplete data, extraction of clinically relevant measurements, and integration with movement analysis.

However, substantial limitations remain, particularly regarding the accuracy achievable with current sensor technology, the challenges of standardizing subject positioning, and the confounding effects of clothing and subject characteristics. Future research directions should focus on enhancing reconstruction accuracy through improved sensor technology and advanced processing algorithms, developing more robust approaches to handling posture variability, and validating derived measurements against gold standard methods across diverse populations.

The integration of 3D body modeling with gait analysis presents particularly promising opportunities for comprehensive assessment of the relationship between body morphology and movement patterns. As these technologies continue to mature, they will likely find increasing application in clinical practice, fitness assessment, and health monitoring, potentially revolutionizing how we understand and address the complex interplay between body composition and physical function.

## 6. Hybrid Systems and Sensor Fusion Strategies for Obesity Detection

Recent advancements in sensor technologies, computational methods, and artificial intelligence have revolutionized approaches to obesity detection and monitoring. This part of the review examines cutting-edge research on hybrid systems and sensor fusion strategies that leverage gait analysis and human voxel modeling for more accurate, non-invasive, and accessible obesity detection. The integration of multiple sensing modalities, privacy-preserving computational approaches, explainable AI methods, and scalable deployment frameworks represents a paradigm shift in obesity management and intervention strategies.

### 6.1. Multimodal / Sensor Fusion System Architectures

Traditional approaches to obesity assessment rely primarily on anthropometric measurements such as BMI, waist circumference, and skinfold thickness. While useful, these methods provide limited insights into the functional implications of excess weight and fail to capture the complex

physiological and biomechanical manifestations of obesity [36]. Gait analysis offers a complementary approach by capturing the biomechanical manifestations of obesity during walking. Early gait analysis systems typically employed single-sensor modalities, such as force plates, optical motion capture, or inertial sensors, each with inherent limitations in comprehensiveness and practical deployment.

The emergence of hybrid systems combining multiple sensor types represents a significant advancement for obesity detection through gait analysis. These multimodal approaches leverage the complementary strengths of different sensing technologies to create more robust, accurate, and practical assessment tools. By simultaneously capturing spatial, temporal, kinematic, and sometimes kinetic parameters of gait, these systems can detect the subtle and complex alterations associated with different degrees of obesity and fat distribution patterns.

Research in this domain has increasingly focused on creating affordable, accessible systems that maintain clinical-grade accuracy [37,38]. This reflects a growing recognition of gait analysis as not merely a research tool but a potential component of routine obesity screening and monitoring programs across various healthcare settings.

#### 6.1.1. Integration of Optical and Depth Sensing Technologies

The integration of RGB cameras with depth sensors has emerged as a foundational approach in obesity detection systems. Kinect-based systems have demonstrated promise by combining RGB imagery with depth information to construct accurate body morphological representations. These systems can generate detailed 3D body models that enable volumetric analysis of body segments—a critical capability for accurate obesity assessment that goes beyond simple anthropometric measurements [37]. The Microsoft Kinect sensor, in particular, has been widely adopted due to its ability to track 25 body joints in real-time while simultaneously capturing RGB and depth data streams, facilitating comprehensive gait and posture analysis in obese individuals.

Recent research demonstrates the potential of combining optical and depth sensing technologies with gait analysis for improved obesity detection and movement assessment. Depth vision sensors, when used alongside wearable sensors, enhance abnormal gait classification [39]. For obese subjects, marker-based optoelectronic systems and wearable magneto-inertial measurement units are commonly used, often integrated with force platforms [40]. An integrated system using depth sensing cameras and IMU sensors, processed through deep learning algorithms, shows significant improvements in gait data accuracy compared to single-method approaches [41]. The Intel RealSense camera, a leading 3D depth sensing technology, has demonstrated promising applications in clinical research, particularly for gait analysis and rehabilitation [42]. These combined technologies offer potential for developing more precise, objective movement-based endpoints for tracking treatment interventions in clinical trials involving obese individuals.

#### 6.1.2. Fusion of Inertial and Optical Sensors

Wearable inertial sensor systems represent another important approach for gait analysis in obesity assessment. These systems typically employ networks of IMUs containing accelerometers, gyroscopes, and sometimes magnetometers attached to various body segments. A wearable magneto-inertial system for gait analysis (H-Gait) has been specifically validated for both normal weight and overweight/obese individuals [43]. This system uses magneto-inertial sensors to capture detailed gait parameters and has demonstrated good reliability across different weight categories.

Inertial sensor systems have shown particular utility for upper limb motion analysis in individuals with obesity, revealing characteristic alterations in arm swing patterns [44]. These systems can quantify parameters such as arm swing amplitude, symmetry, and coordination with lower limb movements—metrics that have proven sensitive to weight-related changes in gait mechanics.

The advantages of wearable inertial systems include portability, ability to function in various environments, and capacity for continuous monitoring during activities of daily living. Recent

miniaturization of IMU technology has led to unobtrusive sensors embedded in clothing, footwear, or accessories, enabling long-term monitoring without significant user burden. However, these systems require careful calibration, synchronization, and drift correction to maintain accuracy.

The combination of IMUs with optical sensing technologies has demonstrated superior performance in characterizing obesity-related gait patterns. IMUs provide detailed information about segment accelerations and orientations, complementing the spatial data captured by optical systems. Research has shown that fusion of these modalities allows for more accurate quantification of gait parameters [45]. At the current state, the presence of bias in the research of Cerfoglio et al. limits the applicability of the inertial-based system in clinics. further research is intended in in this context.

Lee et al. utilized smartphone cameras and wearable IMUs to estimate the knee adduction moment (KAM) and knee flexion moment (KFM), developing a model to optimally diagnose walking patterns and reduce knee load-a particularly relevant application for obese populations who experience greater joint stress during locomotion [37]. The integration of these complementary data streams provided a more holistic understanding of obesity-related biomechanical adaptations than either modality could achieve independently.

### 6.1.3. Thermal Imaging Integration for Multimodal Assessment

Thermal imaging presents a unique opportunity to enhance obesity detection by providing information about subcutaneous fat distribution and brown adipose tissue (BAT) activity. Snekhalatha et al. demonstrated that thermal imaging of abdominal, forearm, and shank regions revealed significant temperature differences between obese and normal-weight individuals, with the abdominal region showing a 4.703% temperature difference [36]. This thermal signature can be attributed to the insulating properties of adipose tissue and altered thermogenesis in obese individuals.

When integrated with skeletal tracking and 3D body modeling, thermal data provides an additional physiological dimension to obesity assessment. Multi-stream architectures that combine thermal, RGB, depth, and inertial data have been proposed, employing various fusion strategies:

1. **Early fusion:** Feature-level integration that combines raw or low-level features from multiple sensors before processing
2. **Late fusion:** Decision-level integration that combines independently processed data from each sensor at the decision stage
3. **Hybrid fusion:** Combinations of early and late fusion approaches that leverage the strengths of each method

Research by Lee et al. introduced a non-contact sensor system that generates 3D body models from 2D images, demonstrating how even limited image inputs (front and side views) can be synthesized into comprehensive 3D body data for obesity monitoring [37]. This approach addresses accessibility issues by reducing hardware requirements while maintaining assessment accuracy.

### 6.1.4. Advanced Data Integration Frameworks

More sophisticated fusion architectures have emerged to handle the heterogeneous data types and sampling rates inherent in multimodal obesity detection systems. Cross-modal attention mechanisms enable systems to dynamically weight the contribution of each modality based on its relevance to specific aspects of obesity assessment. For example, thermal data might receive greater emphasis when evaluating metabolic activity, while inertial and depth data might be prioritized when analyzing gait patterns.

Recent advancements in 3D body model reconstruction have improved upon traditional point cloud techniques. SPLATNet introduces sparse bilateral convolutional layers for efficient point cloud processing, outperforming existing methods in 3D segmentation tasks [46]. Jiang et al. propose a

skeleton-aware approach using PointNet++ and SMPL parameters, incorporating graph aggregation and attention modules for better feature extraction and mapping [47]. Bhatnagar et al. combine implicit function learning with parametric models, using an Implicit Part Network to predict outer and inner body surfaces from sparse point clouds [48]. Their method allows for controllable and accurate 3D reconstructions, even with clothing. Zhou et al. introduce a Gaussian Process layer and adversarial training to encode surface smoothness and shape coherence in their deep autoencoder architecture, demonstrating quantitative improvements over existing DNN-based methods for human body mesh reconstruction from point clouds [49].

## 6.2. Federated Learning and Data Privacy

Gait analysis for obesity detection inherently involves collection of sensitive biometric data that raises significant privacy concerns. Traditional machine learning approaches requiring centralized data aggregation present several problems in this context:

1. **Personal Health Information Protection:** Gait patterns constitute protected health information under regulations like HIPAA and GDPR, necessitating stringent data handling protocols.
2. **Identification Risk:** Gait is a behavioral biometric that can uniquely identify individuals, creating potential for unauthorized tracking or identification if data is compromised.
3. **Stigmatization Concerns:** Data relating to obesity carries social stigma risks, making privacy preservation particularly important for patient dignity and acceptance of monitoring technologies.
4. **Longitudinal Data Vulnerabilities:** Continuous monitoring of gait for obesity management generates extensive personal datasets that, if centralized, create attractive targets for data breaches.

These privacy challenges have historically limited widespread implementation of gait-based obesity monitoring systems, particularly in non-clinical settings like schools or community health programs. The emergence of federated learning approaches offers a promising solution to these concerns by fundamentally changing how models are trained and deployed.

### 6.2.1. Comparative Analysis of FL Algorithms for Obesity Detection

Several federated learning algorithms have been evaluated in the context of gait-based activity recognition, with varying performance characteristics relevant to obesity detection:

1. **Federated Averaging (FedAvg):** The most fundamental FL algorithm works by averaging model updates received from multiple clients before updating the global model. FedAvg performs adequately in homogeneous environments where gait data distributions are similar across users. It offers the advantage of minimizing communication overhead (8.5 MB), making it suitable for resource-constrained devices. However, FedAvg struggles with convergence in heterogeneous settings where gait patterns vary significantly across users with different degrees of obesity [50,51].
2. **Federated Proximal (FedProx):** This extension of FedAvg addresses statistical heterogeneity in federated learning by introducing a proximal term that restricts local model updates, preventing destabilizing changes. We believe that FedProx is particularly valuable for gait-based obesity detection, where individual users may have unique walking patterns influenced

by varying fat distribution, compensatory mechanisms, and comorbidities. By reducing client drift, FedProx ensures more stable learning across diverse populations [50,52].

3. **SCAFFOLD (Stochastic Controlled Averaging for Federated Learning):** This advanced algorithm improves upon both FedAvg and FedProx by incorporating variance reduction techniques. SCAFFOLD corrects for client drift by maintaining control variates that align local model updates with the global model's direction. Comparative studies show SCAFFOLD achieves the highest accuracy (89.1%) and fastest convergence (70 rounds) among FL algorithms for gait analysis. It also demonstrates superior privacy preservation (0.9 privacy score) and explainability (79.4), making it particularly suitable for obesity detection systems that must balance performance with interpretability for clinical use [50].

The selection of an appropriate FL algorithm depends on specific requirements of the obesity detection system, particularly regarding trade-offs between model performance, privacy protection, and deployment constraints. Systems deployed in highly heterogeneous populations (e.g., community-wide screening programs) may benefit from SCAFFOLD's robustness, while resource-constrained applications might prioritize FedAvg's efficiency.

#### 6.2.2. On-Device Learning for Mobile Obesity Screening

On-device learning represents an advanced implementation of federated learning that further enhances privacy and enables real-time obesity risk assessment through gait analysis. This approach performs model training and inference entirely on the user's device, offering several advantages for mobile obesity screening:

1. **Maximum Privacy Protection:** Raw gait data never leaves the device, addressing concerns about collection and storage of sensitive biometric information.
2. **Real-Time Assessment:** Models can provide immediate feedback on obesity-related gait parameters without requiring cloud connectivity, enabling point-of-care applications.
3. **Personalization with Privacy:** Models can adapt to individual walking patterns while still benefiting from population-level insights through federated updates.
4. **Reduced Infrastructure Requirements:** By distributing computational load across user devices, on-device learning reduces need for centralized server infrastructure.

Implementation typically employs lightweight neural networks optimized for mobile processors, with techniques such as model pruning, quantization, and knowledge distillation reducing computational requirements while maintaining accuracy. Research in mobile health applications has demonstrated feasibility of deploying federated learning for health monitoring on resource-constrained devices [53].

The integration of federated learning with Internet of Medical Things architecture has shown promise for obesity risk detection. In these systems, data such as BMI and gait parameters is analyzed locally to assess obesity risk, with expert recommendations generated based on results while preserving user privacy through federated computation [53]. As mobile devices increasingly incorporate advanced sensing capabilities, the potential for widespread, privacy-preserving obesity screening through gait analysis continues to expand.

#### 6.3. Scalable Deployment and Real-Time Systems

The translation of advanced obesity detection technologies from research settings to widespread clinical and community use requires careful consideration of scalability, real-time processing capabilities, and deployment strategies.



### 6.3.1. Edge Computing Architectures for Real-Time Analysis

Real-time obesity detection requires processing complex multimodal data streams with minimal latency. Edge computing architectures that perform computation near the data source rather than in remote data centers have emerged as a preferred approach for these applications.

The custom CNN developed by Snekhalatha et al. for thermal image-based obesity classification was optimized for edge deployment, achieving real-time performance while maintaining high accuracy (92%) [36]. By distributing processing across edge devices and local servers, these systems can deliver immediate feedback during obesity screening sessions without requiring constant connectivity to cloud resources.

Optimization techniques such as model quantization, pruning, and knowledge distillation have been employed to reduce the computational requirements of obesity detection models without sacrificing accuracy. These approaches are particularly important for deployments in resource-constrained settings such as schools and community health centers.

### 6.3.2. School-Based Implementation Strategies

Schools represent critical settings for early obesity detection and intervention. Scalable deployment in educational environments requires systems that are:

1. Non-invasive and respectful of privacy concerns
2. Capable of efficiently screening large numbers of students
3. Simple enough to be operated by school health personnel
4. Affordable within typical school health program budgets

Recent pilot implementations have demonstrated the feasibility of using sensor fusion approaches for school-based obesity screening. These systems typically employ a combination of depth cameras and simplified thermal imaging to assess body composition and movement patterns during physical education activities. The non-contact sensor approach developed by Lee et al. is particularly well-suited for school settings, as it requires minimal equipment and can be integrated into existing health assessment protocols [37].

Privacy considerations are especially important in school implementations, with successful deployments employing federated learning approaches that keep all identifiable data within the school's systems while still benefiting from model improvements across multiple schools.

### 6.3.3. Clinical Integration Frameworks

Integration of advanced obesity detection systems into clinical workflows presents distinct challenges and opportunities. Clinical deployments typically require:

1. Interoperability with existing electronic health record (EHR) systems
2. Compliance with medical device regulations
3. Integration with established clinical assessment protocols
4. Support for longitudinal patient monitoring

Successful clinical implementations have employed modular architectures that separate data acquisition, processing, and visualization components. This approach allows hospitals and clinics to customize deployments based on their specific needs and existing infrastructure.

The thermal imaging approach described by Snekhalatha et al. has been successfully integrated into clinical settings, with the CNN-based classification system achieving an area under the curve (AUC) value of 0.948 in distinguishing obese from normal patients [36]. This performance level makes the system suitable for clinical use as a rapid screening tool, with positive cases referred for more comprehensive assessment.

### 6.3.4. Telemedicine and Remote Monitoring Solutions

The COVID-19 pandemic accelerated the adoption of telemedicine solutions, creating new opportunities for remote obesity monitoring and intervention. Remote monitoring systems typically leverage consumer devices such as smartphones and home cameras to collect data that would previously have required in-person clinical visits.

The approach developed by Lee et al., which generates 3D body models from simple 2D images, is particularly well-suited for telemedicine applications [37]. Patients can capture front and side images using their smartphones, with the system generating detailed body composition analyses that can be reviewed by healthcare providers during virtual consultations.

These remote monitoring solutions employ several strategies to ensure data quality and reliability:

1. Standardized capture protocols with real-time guidance
2. Automated quality control to reject unsuitable images
3. Calibration procedures to account for varying camera characteristics
4. Confidence metrics that indicate measurement reliability

The integration of these systems with telehealth platforms creates comprehensive obesity management solutions that combine detection, monitoring, and intervention components within unified user experiences.

To sum up, the integration of hybrid systems and sensor fusion strategies for obesity detection represents a significant advancement over traditional assessment methods. By combining multiple sensing modalities-including optical, depth, inertial, and thermal technologies-these systems provide more comprehensive and accurate characterizations of obesity-related physiological and biomechanical alterations. The incorporation of federated learning approaches addresses critical privacy concerns while enabling continuous model improvement, while explainable AI techniques translate complex sensor data into clinically actionable insights. Scalable deployment architectures facilitate the implementation of these technologies across diverse settings, from schools to clinics to home environments, creating new opportunities for early intervention and ongoing management of obesity.

Future research directions should focus on further integration of metabolic and behavioral sensing modalities, refinement of privacy-preserving learning techniques, development of more intuitive explanatory frameworks, and validation of these systems in diverse real-world settings. As these technologies mature, they have the potential to transform obesity detection and management from periodic clinical assessments to continuous, personalized monitoring and intervention.

## 7. Future Directions and Research Opportunities for Obesity Detection Based on Gait Analysis

Recent advancements in gait analysis technologies have opened new avenues for obesity detection and management. This section explores emerging research directions that promise to enhance the accuracy, accessibility, and clinical utility of gait-based obesity screening tools while addressing current methodological limitations.

### 7.1. Toward Portable, AI-Enabled Obesity Detection

The development of compact, low-power screening tools represents a critical frontier in democratizing obesity detection. Smartphone-based gait analytics have demonstrated particular promise, leveraging ubiquitous devices to capture inertial measurement unit (IMU) data from built-in accelerometers and gyroscopes. A 2024 study achieved 97% classification accuracy for obesity using a hybrid CNN-LSTM model trained on smartphone-derived gait patterns from 138 subjects [6]. These systems eliminate the need for specialized laboratory equipment, enabling large-scale screening in community settings.

Low-cost wearable insoles with embedded pressure sensors further enhance portability while providing direct measurements of ground reaction forces. Early prototypes demonstrated the feasibility of Bluetooth-enabled smartshoes for gait monitoring [54], though recent innovations in sensor miniaturization and energy efficiency have improved their practicality for continuous use. Integration of edge computing architectures allows real-time gait analysis without reliance on cloud infrastructure, addressing privacy concerns and latency issues.

### 7.2. Standardized Protocols and Open Datasets

Current research suffers from fragmented methodologies, as evidenced by a 2024 meta-analysis identifying significant heterogeneity in gait parameter reporting across 14 obesity studies [55]. Establishing annotated obesity gait libraries with ground-truth validation requires multidisciplinary collaboration to define:

1. Unified spatiotemporal parameter definitions
2. Standardized BMI classification thresholds
3. Age- and sex-specific normative ranges

The Health&Gait dataset represents a pioneering effort in this direction, comprising 1,564 video samples from 398 participants with synchronized anthropometric and gait data [56]. However, critical gaps persist in pediatric populations, where obesity-induced gait modifications differ substantially from adults. A 2025 intervention study in obese children highlighted the need for youth-specific benchmarks, demonstrating unique pelvic kinematic adaptations during walking [13].

Open challenges include reconciling optical motion capture with wearable sensor outputs and developing cross-modal calibration protocols. Shared benchmarks must account for ethnic diversity, socioeconomic factors, and comorbid conditions to avoid algorithmic bias in heterogeneous populations.

### 7.3. Wearable and Optical Sensor Integration

Multimodal sensor fusion approaches are overcoming the limitations of single-modality systems. The INDIP platform exemplifies this trend, combining plantar pressure insoles, inertial measurement units (IMUs), and time-of-flight distance sensors to achieve  $\leq 0.06$  m stride length error across diverse cohorts including Parkinson's and COPD patients [57]. Integrating camera-derived kinematic data with wearable heart rate (HR) and IMU metrics enables holistic health monitoring – a concept validated in video-based systems achieving 94% sex classification accuracy using gait features [56].

Emerging technologies leverage computer vision to extract 3D joint kinematics from smartphone videos, bypassing the need for marker-based systems. When combined with wearable-derived cardiovascular metrics, these systems can correlate gait abnormalities with metabolic parameters like VO<sub>2</sub> max. However, lighting variability and occlusion remain technical hurdles, necessitating advanced neural networks trained on augmented datasets simulating real-world conditions.

### 7.4. Personalization with Digital Twins

Patient-specific digital twin models are revolutionizing intervention planning by simulating gait adaptations underweight change scenarios. A 2025 kinetic study demonstrated the predictive value of such models, showing improved pelvic kinematics in obese children following six-month exercise programs [13]. These virtual replicas integrate:

- Biomechanical body composition profiles
- Muscle activation patterns
- Joint loading characteristics

Deep learning architectures trained on longitudinal gait data can forecast individualized responses to dietary, surgical, or exercise interventions. For instance, transformer-based models show

promise in predicting post-bariatric surgery gait normalization trajectories using preoperative spatiotemporal parameters [55]. Federated learning frameworks enable model refinement across institutions while maintaining data privacy – a crucial consideration for sensitive health data.

The convergence of wearable technologies, advanced analytics, and personalized modeling heralds a new era in obesity detection and management. Realizing this potential requires sustained investment in standardized datasets, interoperable sensor platforms, and validation studies across diverse populations. Priorities include expanding pediatric gait databases, developing ethical AI governance frameworks, and translating laboratory innovations into scalable public health solutions. By addressing these challenges, gait analysis may soon become a cornerstone of precision medicine approaches to obesity [6,13,54–58].

## 8. Conclusions

The advancement of optical sensor-based gait systems represents a significant leap in the field of motion analysis, with relevance to the diagnosis and monitoring of obesity and related gait dysfunctions. This review has explored the evolving landscape of gait sensing technologies, particularly emphasizing non-invasive, markerless, and hybrid optical systems that combine visual, depth, and inertial sensing modalities. Through a multidisciplinary lens encompassing biomedical engineering, computer vision, and clinical sciences, the article has illustrated how such systems contribute to more accessible, accurate, and context-sensitive evaluations of human locomotion.

One of the core contributions of this review lies in its synthesis of the technological evolution of optical gait sensing tools. Beginning with traditional marker-based motion capture systems, such as Vicon and OptiTrack, and progressing toward markerless alternatives employing deep learning and video-based pose estimation (e.g., OpenPose, MediaPipe, and DeepLabCut), the review charts a clear trajectory of innovation aimed at reducing invasiveness and increasing ecological validity. Markerless video-based systems have opened new avenues for gait assessment in naturalistic settings, making them especially valuable for longitudinal monitoring and large-scale screening.

A key theme throughout the review is the transition from laboratory-bound assessment to portable, real-world deployment. This shift is exemplified by the emergence of low-cost RGB-D cameras like the Microsoft Kinect and Intel RealSense series, which enable three-dimensional tracking of body movements without the need for specialized facilities or technical operators. When integrated with machine learning models and advanced biomechanical analysis pipelines, these systems demonstrate strong potential for clinical translation, particularly in resource-constrained environments. The feasibility of such systems in pediatric populations affected by obesity—where early diagnosis and intervention are crucial—is an especially compelling application scenario that merits continued attention.

Another notable insight concerns the proliferation of hybrid sensor systems, which combine optical sensors with wearable inertial measurement units (IMUs), force plates, or bioimpedance analyzers. These multimodal systems provide richer datasets by fusing kinematic, kinetic, and physiological information, thus supporting a more holistic understanding of gait biomechanics. For instance, the integration of gait data with body composition analysis has the potential to uncover correlations between specific biomechanical patterns and underlying adiposity-related dysfunctions. Such systems not only improve diagnostic precision but also pave the way for personalized intervention strategies.

The review also underscores the critical role of artificial intelligence, particularly deep learning and computer vision, in advancing the capabilities of optical gait analysis. Convolutional neural networks (CNNs), graph convolutional networks (GCNs), and recurrent architectures like LSTMs have been employed to enhance the robustness of pose estimation, detect anomalies, and classify gait types. These models often outperform traditional rule-based algorithms, especially in unconstrained environments where occlusions, lighting variations, and individual variability present significant challenges. Nevertheless, the review also calls for transparency in AI models—emphasizing explainability and fairness as essential components of clinical-grade technologies.

Despite these promising developments, several limitations and challenges remain. Chief among them is the issue of validation and standardization. While numerous studies demonstrate high accuracy under controlled conditions, real-world deployment often exposes vulnerabilities in sensor calibration, tracking fidelity, and generalizability across populations. This is particularly pertinent in the context of pediatric obesity, where age-specific gait patterns, compliance issues, and varying body morphologies introduce further complexities. Addressing these challenges will require harmonized benchmarking protocols, open datasets representing diverse demographics, and interdisciplinary collaboration between engineers, clinicians, and public health stakeholders.

Another area warranting further exploration is the ethical and privacy implications of video-based monitoring. Markerless gait analysis often involves the capture and processing of sensitive visual data, raising concerns around informed consent, data anonymization, and long-term storage. Ethical frameworks must evolve in parallel with technical capabilities to ensure responsible deployment, especially in vulnerable populations such as children or individuals with disabilities. Solutions such as edge computing, federated learning, and privacy-preserving AI offer promising pathways to mitigate these risks while preserving data utility.

In terms of clinical impact, the review highlights how optical gait systems can support early screening, risk stratification, and intervention planning for obesity and associated comorbidities. By capturing subtle biomechanical deviations in everyday contexts, these systems can aid in identifying individuals at risk before overt symptoms emerge. Moreover, the non-contact nature of optical systems enhances user comfort and adherence, facilitating more frequent and longitudinal assessments that are critical for monitoring treatment efficacy. These attributes align with the broader shift toward preventive, personalized, and participatory (P4) healthcare models.

From a research perspective, several future directions emerge. First, the integration of gait analysis with other health indicators—such as metabolic biomarkers, cardiovascular metrics, or psychological parameters—can offer more comprehensive phenotyping of obesity-related dysfunctions. Second, advances in real-time data processing and edge AI could enable closed-loop feedback systems for rehabilitation or physical activity interventions, thereby transforming gait monitoring from a diagnostic to a therapeutic tool. Third, the development of explainable and ethically aligned AI models remains a priority, particularly in clinical contexts where trust, accountability, and interpretability are paramount.

Furthermore, to ensure broad adoption, there is a need for user-centered design approaches that consider the perspectives of end-users, including patients, caregivers, and clinicians. Systems must be intuitive, affordable, and seamlessly integrable into existing workflows. Collaborative efforts across academia, industry, and healthcare systems will be essential to co-create solutions that are both technically sound and socially acceptable. Publicly funded initiatives and open-source ecosystems can play a pivotal role in democratizing access and accelerating innovation.

In conclusion, optical sensor-based gait systems have matured from niche laboratory tools into versatile platforms capable of addressing real-world clinical challenges, particularly in the context of obesity detection and gait dysfunction analysis. Their non-invasive, scalable, and increasingly intelligent nature positions them as integral components of next-generation healthcare and biomechanical research. However, their full potential will only be realized through rigorous validation, ethical deployment, and sustained interdisciplinary collaboration. By aligning technological advancement with clinical need and societal values, these systems hold promise not only for advancing gait science but also for improving human health across the lifespan.

## References

- a. One in Eight People Are Now Living with Obesity Available online: <https://www.who.int/news/item/01-03-2024-one-in-eight-people-are-now-living-with-obesity> (accessed on 18 April 2025).
- b. Obesity and Overweight Available online: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> (accessed on 18 April 2025).
2. World\_Obesity\_Atlas\_2023\_Report.Pdf.

3. Obesity | What We Do. *World Heart Federation*.
4. Koinis, L.; Maharaj, M.; Natarajan, P.; Fonseka, R.D.; Fernando, V.; Mobbs, R.J. Exploring the Influence of BMI on Gait Metrics: A Comprehensive Analysis of Spatiotemporal Parameters and Stability Indicators.
5. Degbey, G.-S.; Hwang, E.; Park, J.; Lee, S. Deep Learning-Based Obesity Identification System for Young Adults Using Smartphone Inertial Measurements. *International Journal of Environmental Research and Public Health* **2024**, *21*, 1178, doi:10.3390/ijerph21091178.
6. Muro-de-la-Herran, A.; Garcia-Zapirain, B.; Mendez-Zorrilla, A. Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications. *Sensors (Basel)* **2014**, *14*, 3362–3394, doi:10.3390/s140203362.
7. Das, R.; Paul, S.; Mourya, G.K.; Kumar, N.; Hussain, M. Recent Trends and Practices Toward Assessment and Rehabilitation of Neurodegenerative Disorders: Insights From Human Gait. *Front. Neurosci.* **2022**, *16*, doi:10.3389/fnins.2022.859298.
8. Carbajales-Lopez, J.; Becerro-de-Bengoa-Vallejo, R.; Losa-Iglesias, M.E.; Casado-Hernández, I.; Benito-De Pedro, M.; Rodríguez-Sanz, D.; Calvo-Lobo, C.; San Antolín, M. The OptoGait Motion Analysis System for Clinical Assessment of 2D Spatio-Temporal Gait Parameters in Young Adults: A Reliability and Repeatability Observational Study. *Applied Sciences* **2020**, *10*, 3726, doi:10.3390/app10113726.
9. Naz, A.; Prasad, P.; McCall, S.; Chan, C.L.; Ochi, I.; Gong, L.; Yu, M. Privacy-Preserving Abnormal Gait Detection Using Computer Vision and Machine Learning.
10. Desrochers, P.C.; Kim, D.; Keegan, L.; Gill, S.V. Association between the Functional Gait Assessment and Spatiotemporal Gait Parameters in Individuals with Obesity Compared to Normal Weight Controls: A Proof-of-Concept Study. *J Musculoskelet Neuronal Interact* **2021**, *21*, 335–342.
11. Tao, W.; Liu, T.; Zheng, R.; Feng, H. Gait Analysis Using Wearable Sensors. *Sensors (Basel)* **2012**, *12*, 2255–2283, doi:10.3390/s120202255.
12. Popescu, C.; Matei, D.; Amzolini, A.M.; Trăistaru, M.R. Comprehensive Gait Analysis and Kinetic Intervention for Overweight and Obese Children. *Children* **2025**, *12*, 122, doi:10.3390/children12020122.
13. Prisco, G.; Pirozzi, M.A.; Santone, A.; Esposito, F.; Cesarelli, M.; Amato, F.; Donisi, L. Validity of Wearable Inertial Sensors for Gait Analysis: A Systematic Review. *Diagnostics (Basel)* **2024**, *15*, 36, doi:10.3390/diagnostics15010036.
14. Apovian, C.M. Obesity: Definition, Comorbidities, Causes, and Burden. *Am J Manag Care* **2016**, *22*, s176–185.
15. Liang, S.; Zhang, Y.; Diao, Y.; Li, G.; Zhao, G. The Reliability and Validity of Gait Analysis System Using 3D Markerless Pose Estimation Algorithms. *Front Bioeng Biotechnol* **2022**, *10*, 857975, doi:10.3389/fbioe.2022.857975.
16. Nakano, N.; Sakura, T.; Ueda, K.; Omura, L.; Kimura, A.; Iino, Y.; Fukashiro, S.; Yoshioka, S. Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Front. Sports Act. Living* **2020**, *2*, doi:10.3389/fspor.2020.00050.
17. MediaPipe Pose — DroneVis 1.3.0 Documentation Available online: <https://dronevis.readthedocs.io/en/latest/pose/mediapipe.html> (accessed on 1 May 2025).
18. MediaPipe Pose Estimation · Models · Dataloop Available online: [https://dataloop.ai/library/model/qualcomm\\_mediapipe-pose-estimation/](https://dataloop.ai/library/model/qualcomm_mediapipe-pose-estimation/) (accessed on 1 May 2025).
19. Mathis, A.; Mamidanna, P.; Cury, K.M.; Abe, T.; Murthy, V.N.; Mathis, M.W.; Bethge, M. DeepLabCut: Markerless Pose Estimation of User-Defined Body Parts with Deep Learning. *Nat Neurosci* **2018**, *21*, 1281–1289, doi:10.1038/s41593-018-0209-y.
20. Insafutdinov, E.; Pishchulin, L.; Andres, B.; Andriluka, M.; Schiele, B. DeeperCut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model. In Proceedings of the Computer Vision – ECCV 2016; Leibe, B., Matas, J., Sebe, N., Welling, M., Eds.; Springer International Publishing: Cham, 2016; pp. 34–50.
21. Lauer, J.; Zhou, M.; Ye, S.; Menegas, W.; Nath, T.; Rahman, M.M.; Santo, V.D.; Soberanes, D.; Feng, G.; Murthy, V.N.; et al. Multi-Animal Pose Estimation and Tracking with DeepLabCut 2021, 2021.04.30.442096.
22. Panconi, G.; Grasso, S.; Guarducci, S.; Mucchi, L.; Minciacci, D.; Bravi, R. DeepLabCut Custom-Trained Model and the Refinement Function for Gait Analysis. *Sci Rep* **2025**, *15*, 2364, doi:10.1038/s41598-025-85591-1.

23. From Marker to Markerless: Validating DeepLabCut for 2D Sagittal Plane Gait Analysis in Adults and Newly Walking Toddlers - ScienceDirect Available online: <https://www.sciencedirect.com/science/article/pii/S0021929025002209> (accessed on 1 May 2025).
24. Albert, J.A.; Owolabi, V.; Gebel, A.; Brahms, C.M.; Granacher, U.; Arnrich, B. Evaluation of the Pose Tracking Performance of the Azure Kinect and Kinect v2 for Gait Analysis in Comparison with a Gold Standard: A Pilot Study. *Sensors (Basel)* **2020**, *20*, 5104, doi:10.3390/s20185104.
25. The Development and Evaluation of a Fully Automated Markerless Motion Capture Workflow. *Journal of Biomechanics* **2022**, *144*, 111338, doi:10.1016/j.jbiomech.2022.111338.
26. Zhang, C.; Greve, C.; Verkerke, G.J.; Roossien, C.C.; Houdijk, H.; Hijmans, J.M. Pilot Validation Study of Inertial Measurement Units and Markerless Methods for 3D Neck and Trunk Kinematics during a Simulated Surgery Task. *Sensors (Basel)* **2022**, *22*, 8342, doi:10.3390/s22218342.
27. Comparison of Depth Cameras for Three-Dimensional Reconstruction in Medicine - Chuang-Yuan Chiu, Michael Thelwell, Terry Senior, Simon Choppin, John Hart, Jon Wheat, 2019 Available online: <https://journals.sagepub.com/doi/abs/10.1177/0954411919859922> (accessed on 4 May 2025).
28. Weiss, A.; Hirshberg, D.; Black, M.J. Home 3D Body Scans from Noisy Image and Range Data. In Proceedings of the 2011 International Conference on Computer Vision; November 2011; pp. 1951–1958.
29. Li, Z.; Oskarsson, M.; Heyden, A. Detailed 3D Human Body Reconstruction from Multi-View Images Combining Voxel Super-Resolution and Learned Implicit Representation. *Appl Intell* **2022**, *52*, 6739–6759, doi:10.1007/s10489-021-02783-8.
30. Ruget, A.; Tyler, M.; Mora Martín, G.; Scholes, S.; Zhu, F.; Gyongy, I.; Hearn, B.; McLaughlin, S.; Halimi, A.; Leach, J. Pixels2Pose: Super-Resolution Time-of-Flight Imaging for 3D Pose Estimation. *Science Advances* **2022**, *8*, eade0123, doi:10.1126/sciadv.ade0123.
31. Tsoli, A.; Loper, M.; Black, M.J. Model-Based Anthropometry: Predicting Measurements from 3D Human Scans in Multiple Poses. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision; March 2014; pp. 83–90.
32. Alexa, A. Measurement Methods for Body Fat (%) Assessment from 3D Kinect Scans. **2016**.
33. A Multidomain Approach to Assessing the Convergent and Concurrent Validity of a Mobile Application When Compared to Conventional Methods of Determining Body Composition Available online: <https://www.mdpi.com/1424-8220/20/21/6165> (accessed on 4 May 2025).
34. Feature Hiding in 3D Human Body Scans\* - Joseph Laws, Nathaniel Bauernfeind, Yang Cai, 2006 Available online: <https://journals.sagepub.com/doi/10.1057/palgrave.ivs.9500136?icid=int.sj-abstract.similar-articles.7> (accessed on 4 May 2025).
35. Computer Aided Diagnosis of Obesity Based on Thermal Imaging Using Various Convolutional Neural Networks - ScienceDirect Available online: <https://www.sciencedirect.com/science/article/abs/pii/S1746809420303633> (accessed on 4 May 2025).
36. Lee, J.-Y.; Kwon, K.; Kim, C.; Youm, S. Development of a Non-Contact Sensor System for Converting 2D Images into 3D Body Data: A Deep Learning Approach to Monitor Obesity and Body Shape in Individuals in Their 20s and 30s. *Sensors* **2024**, *24*, 270, doi:10.3390/s24010270.
37. Ergün, U.; Aktepe, E.; Koca, Y.B. Detection of Body Shape Changes in Obesity Monitoring Using Image Processing Techniques. *Sci Rep* **2024**, *14*, 24178, doi:10.1038/s41598-024-73270-6.
38. Wong, C.; McKeague, S.; Correa, J.; Liu, J.; Yang, G.-Z. Enhanced Classification of Abnormal Gait Using BSN and Depth. In Proceedings of the 2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks; May 2012; pp. 166–171.
39. Monfrini, R.; Rossetto, G.; Scalona, E.; Galli, M.; Cimolin, V.; Lopomo, N.F. Technological Solutions for Human Movement Analysis in Obese Subjects: A Systematic Review. *Sensors (Basel)* **2023**, *23*, 3175, doi:10.3390/s23063175.
40. Bersamira, J.N.; De Chavez, R.J.A.; Salgado, D.D.S.; Sumilang, M.M.C.; Valles, E.R.; Roxas, E.A.; dela Cruz, A.R. Human Gait Kinematic Estimation Based on Joint Data Acquisition and Analysis from IMU and Depth-Sensing Camera. In Proceedings of the 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM); November 2019; pp. 1–6.

41. Siena, F.L.; Byrom, B.; Watts, P.; Breedon, P. Utilising the Intel RealSense Camera for Measuring Health Outcomes in Clinical Research. *J Med Syst* **2018**, *42*, 53, doi:10.1007/s10916-018-0905-x.
42. Agostini, V.; Gastaldi, L.; Rosso, V.; Knaflitz, M.; Tadano, S. A Wearable Magneto-Inertial System for Gait Analysis (H-Gait): Validation on Normal Weight and Overweight/Obese Young Healthy Adults. *Sensors* **2017**, *17*, 2406, doi:10.3390/s17102406.
43. Use of Inertial Sensor System for Upper Limb Motion Analysis in Obese Subjects: Preliminary Setting and Analysis. *Gait & Posture* **2022**, *97*, S124–S125, doi:10.1016/j.gaitpost.2022.07.086.
44. Cerfoglio, S.; Lopomo, N.F.; Capodaglio, P.; Scalona, E.; Monfrini, R.; Verme, F.; Galli, M.; Cimolin, V. Assessment of an IMU-Based Experimental Set-Up for Upper Limb Motion in Obese Subjects. *Sensors* **2023**, *23*, 9264, doi:10.3390/s23229264.
45. Su, H.; Jampani, V.; Sun, D.; Maji, S.; Kalogerakis, E.; Yang, M.-H.; Kautz, J. SPLATNet: Sparse Lattice Networks for Point Cloud Processing. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; IEEE: Salt Lake City, UT, June 2018; pp. 2530–2539.
46. Jiang, H.; Cai, J.; Zheng, J. Skeleton-Aware 3D Human Shape Reconstruction From Point Clouds. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV); October 2019; pp. 5430–5440.
47. Bhatnagar, B.L.; Sminchisescu, C.; Theobalt, C.; Pons-Moll, G. Combining Implicit Function Learning and Parametric Models for 3D Human Reconstruction. In Proceedings of the Computer Vision – ECCV 2020; Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M., Eds.; Springer International Publishing: Cham, 2020; pp. 311–329.
48. Zhou, B.; Franco, J.-S.; Bogu, F.; Tekin, B.; Boyer, E. Reconstructing Human Body Mesh from Point Clouds by Adversarial GP Network. In Proceedings of the Computer Vision – ACCV 2020; Ishikawa, H., Liu, C.-L., Pajdla, T., Shi, J., Eds.; Springer International Publishing: Cham, 2021; pp. 123–139.
49. Prathusha, D.P.; Aparna, K. A Comparative Analysis of FedAvg, FedProx, and Scaffold in Gait-Based Activity Recognition by Evaluating Accuracy, Privacy, and Explainability. *Global Journal of Engineering Innovations and Interdisciplinary Research* **2025**, *5*, doi:10.33425/3066-1226.1067.
50. Wang, H.; Yurochkin, M.; Sun, Y.; Papailiopoulos, D.; Khazaeni, Y. Federated Learning with Matched Averaging 2020.
51. Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated Optimization in Heterogeneous Networks. *Proceedings of Machine Learning and Systems* **2020**, *2*, 429–450.
52. Wang, T.; Du, Y.; Gong, Y.; Choo, K.-K.R.; Guo, Y. Applications of Federated Learning in Mobile Health: Scoping Review. *J Med Internet Res* **2023**, *25*, e43006, doi:10.2196/43006.
53. TOWARDS A LOW POWER WIRELESS SMARTSHOE SYSTEM FOR GAIT ANALYSIS IN PEOPLE WITH DISABILITIES Available online: <https://www.resna.org/sites/default/files/conference/2015/cac/zerin.html> (accessed on 4 May 2025).
54. Scataglini, S.; Dellaert, L.; Meeuwssen, L.; Staeljanssens, E.; Truijien, S. The Difference in Gait Pattern between Adults with Obesity and Adults with a Normal Weight, Assessed with 3D-4D Gait Analysis Devices: A Systematic Review and Meta-Analysis. *Int J Obes (Lond)* **2025**, *49*, 541–553, doi:10.1038/s41366-024-01659-4.
55. Zafra-Palma, J.; Marín-Jiménez, N.; Castro-Piñero, J.; Cuenca-García, M.; Muñoz-Salinas, R.; Marín-Jiménez, M.J. Health & Gait: A Dataset for Gait-Based Analysis. *Sci Data* **2025**, *12*, 44, doi:10.1038/s41597-024-04327-4.
56. Frontiers | A Multi-Sensor Wearable System for the Assessment of Diseased Gait in Real-World Conditions Available online: <https://www.frontiersin.org/journals/bioengineering-and-biotechnology/articles/10.3389/fbioe.2023.1143248/full> (accessed on 4 May 2025).
57. Ahmed, U.; Ali, M.F.; Javed, K.; Babri, H.A. Predicting Physiological Developments from Human Gait Using Smartphone Sensor Data 2017.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s)



disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.