Journal of Applied Engineering and Technological Science
Vol 7(2) 2026: 1703-1718

A NON-CONTACT HEIGHT MEASUREMENT METHOD USING MEDIAPIPE AND YOLOV8 IN A THREE-DIMENSION SPACE

Phu Nguyen Trung[1*], Dat Quoc Dong[2], Kien Nguyen Phan[3*]
Hanoi University of Industry, Vietnam[1]
Hanoi University of Science and Technology, Vietnam[2,3]
phunt@haui.edu.vn[1], kien.nguyenphan@hust.edu.vn[3]

Received: 23 July 2025, Revised: 06 February 2026, Accepted: 27 February 2026
**Corresponding Author*

ABSTRACT
Accurate human height measurement is an essential clinical parameter for calculating body mass index and assessing patient health status. However, traditional contact-based or manual methods are impractical for patients with limited mobility or during emergency conditions. This study addresses this limitation by developing a non-contact 3D height estimation system using a smartphone-based vision approach. The system employs MediaPipe to extract skeletal landmarks and a regression model to predict human height from bone segment lengths and a reference object. Experimental evaluation on 166 subjects across multiple positions demonstrates a high accuracy with an average error rate of 1.54 ± 0.64%, outperforming most existing camera-based systems. The proposed method offers a practical, low-cost, and portable solution for medical height assessment, particularly beneficial in clinical and emergency settings. Potential limitations include dependence on camera angle and lighting conditions, which will be addressed in future work.
Keywords : *Non-Contact, Height Measurement, 3D Space, Mediapipe*

1. Introduction

Body mass index (BMI) has been used as an important metric to evaluate development in children, assess the risk for certain diseases such as diabetes and hypertension (Khanna et al., 2022; Zhang et al., 2025; Efeti et al., 2025; Nayak S S et al., 2024), and determine dose-related metrics such as radiation dose in abdominal computed tomography (O'Neill et al., 2018) and general radiography (Lee et al., 2020). To calculate BMI, information about weight and height is needed. In healthy individuals, measurement of weight and height is straightforward. Weight is often measured with a scale, while height is commonly measured with a tape measure, height gauge, or stadiometer. However, most of these tools require a person to stand upright, which is not feasible for patients with limited mobility. There are chair scales that allow patients to sit in a wheelchair while measuring their weight. Additionally, developed a bedside scale that can measure patients' weight in a lying position (Duc et al., 2023; Eriks-Hoogland et al., 2011). While it is not practical and perhaps not accurate to measure patients' height in a sitting position, it is possible to measure their height in a lying position. However, these situations usually require additional time and staff to straighten the patient's body for an accurate measurement. Additionally, height measurement often requires direct patient contact, which can be dangerous for patients with highly contagious infectious diseases. In fact, there have been many studies measuring human body height, but measuring the height of intensive care patients or patients who are unable to move is extremely limited. While this height measurement is closely related to the treatment parameters for patients (Nirmal et al., 2015; Mikula et al., 2016; Dennis et al., 2015). There are even studies that show the ability of medical staff in the intensive care unit to estimate height by eye with patients, but the results of these studies do not have much meaning in determining the overall accuracy of all other doctors (Venkataraman, et al., 2015) (Ahmad et al., 2025; L'her et al., 2016). For example, during the recent Covid-19 period, many hospitals in Vietnam lacked intensive care doctors and forced the Vietnamese Ministry of Health to mobilize doctors in related specialties to support patient care. In addition to not ensuring expertise, height determination cannot be consistent with the ability to determine the height of patients in intensive care by doctors or medical staff in the intensive care unit. This highlights a critical research gap in developing an automated, non-contact height measurement method applicable to emergency

and intensive care contexts, where manual procedures are unreliable and potentially unsafe. Recent studies (2019–2024) have emphasized the need for vision-based, AI-driven systems that minimize physical contact while maintaining clinical accuracy. Therefore, developing a non-contact height measurement system for critically ill patients, who are unconscious or have lost mobility, has become a significant research goal. Recent advances in non-contact imaging have explored various technologies, including LiDAR sensors, stereo vision, and depth cameras, for anthropometric measurement. LiDAR-based approaches offer high accuracy but remain costly and impractical for clinical deployment, while stereo-vision systems are sensitive to lighting and require complex calibration. These limitations underscore the need for a cost-effective, camera-based system deployable in real-world healthcare environments (Krzeszowski et al., 2023; Wang et al., 2024).

Non-contact height measurement has primarily been performed using image processing and artificial intelligence (Chang et al., 2024; Ismail et al., 2020; Kim et al., 2023; Chai & Cao, 2018; Li et al., 2015; Dokthurian et al., 2021; Ciampini et al., 2024; Tonini et al., 2022; Velesaca et al., 2023) developed a real-time height measurement algorithm for video surveillance). After correcting for camera distortion and extracting the foreground, they obtained information about a person's head vertices and foot points. This information was then combined with the height of two reference objects (a wall and a door frame) to derive human height. Although the reported error was low (the highest error was 1.2 cm), the study was conducted with a small population (n = 6 subjects). Additionally, the use of hats and high-heeled shoes can hinder the accuracy of extracting head vertices and foot points. Haritosh used convolutional neural networks to extract facial features and artificial neural networks to estimate body height (Haritosh et al., 2019). Their method resulted in an error rate of approximately 7.3 cm, which might be impractical for determining medical treatment. In another approach, researchers have applied a convolutional neural network model to human images captured by a depth camera to measure human height (Lee et al., 2020; Trivedi et al., 2021; Yin & Zhou, 2020). Although they were able to estimate human height with a low error rate, the use of a depth camera may limit its practicality. Additionally, the accuracy of their method can be affected by factors such as camera focal length, camera angle, lighting conditions in the surrounding environment, and the use of hats or shoes. Quantitative benchmarks from prior studies indicate that most camera-based or AI-assisted methods report errors ranging from 1.2 to 7.3 cm, depending on sensor type and environmental conditions. However, few studies provide clinical validation, making it difficult to assess their applicability to emergency or critical care scenarios.

Our previous study proposed the use of MediaPipe to extract the coordinates of skeletal key points from a human image and apply a multiple linear regression model to estimate human height, which helps improve the accuracy of height measurement (Nguyen Trung et al., 2024; Phan et al., 2023; Kien, Van Thao, et al., 2025; Kien, Dat, et al., 2025; Phan et al., 2024; Dill et al., 2024; Anh et al., 2025). However, in that study, human height was evaluated in 2D space. Because 2D space only captures the vertical dimension of a person in a flat plane, 2D height measurement is susceptible to head and body tilt. On the other hand, 3D space captures both the vertical dimension and depth, providing a more complete picture of the person's height. Using the dataset collected from our previous study (Nguyen Trung et al., 2024), this study developed a non-contact height measurement method in 3D space. In addition to the four positions—standing upright, 45-degree rotation, 90-degree rotation, and kneeling—described in our previous study, this study adds two additional simulated lying positions, including supine and supine with bent knees. Unlike previous implementations, this work introduces the integration of MediaPipe (Lugaresi et al., 2019; Dill et al., 2024; Kappan et al., 2025) with YOLOv8 (Terven et al., 2023; Kothari & Chakurkar, 2025) for joint detection and reference object localization, establishing a unified pipeline for 3D human height estimation (Lai et al., 2023). This fusion allows accurate extraction of skeletal key points and spatial calibration, addressing challenges of body tilt and variable viewing angles. We hypothesize that 3D height measurement will significantly enhance the accuracy of estimation in complicated positions, such as bent knees, compared to other straight-leg positions. While most existing methods rely on either 2D image estimation or expensive depth sensors, our approach introduces several scientific contributions compared with state-of-the-art techniques (Lai et al., 2023). Specifically, we employ a standard smartphone

camera combined with a physical reference object, eliminating the need for depth sensors and ensuring flexibility in real clinical settings a 3D geometric correction is proposed to compensate for the Z-axis distortion inherent in MediaPipe's depth estimation, which allows more accurate reconstruction of upright human posture. An automatic height estimation algorithm is then developed using the corrected 3D coordinates, overcoming the limitations of conventional 2D regression models. In theoretical terms, the proposed system models the geometric distortion between projected and true skeletal coordinates through a regression-based correction matrix, enhancing spatial reconstruction accuracy. This analytical formulation distinguishes our method from previous empirical-only studies. These contributions not only improve the measurement accuracy under complex postures (such as bent knees or partial supine positions) but also demonstrate the feasibility of applying 3D correction to smartphone-based anthropometric systems without additional hardware. In particular, the integration of YOLOv8 (Juan Terven, 2023)for precise reference-object localization with MediaPipe-based skeletal keypoint extraction establishes a unified 3D estimation pipeline—representing a methodological advancement beyond previous 2D regression-based systems (Nguyen Trung et al., 2024; Phan et al., 2023; Kien, Van Thao, et al., 2025; Kien, Dat, et al., 2025; Phan et al., 2024). This integration enables automatic calibration of pixel-to-real-world scaling, significantly reducing spatial distortion and enhancing robustness under varying illumination and camera angles. Compared to earlier depth-camera and stereo-vision methods, which typically report errors between 1.2 cm and 7.3 cm (Chai & Cao, 2018; Haritosh et al., 2019; Lee et al., 2020; Trivedi et al., 2021; Yin & Zhou, 2020) the proposed framework achieves comparable or better accuracy while relying solely on a standard smartphone camera, highlighting both its technical efficiency and clinical practicality. This highlights both the scientific novelty and the practical significance of the proposed method compared with previous works. From a theoretical perspective, the study introduces a geometric correction matrix that refines MediaPipe's Z-axis depth inference, thus improving the mathematical mapping from 2D to 3D joint coordinates. This analytical formulation provides a foundation for future non-contact anthropometric estimation models. By utilizing OpenCV (Bradski, 2000) and MediaPipe (Lugaresi et al., 2019), the third spatial dimension of an image is defined based on the depth information extracted from the image. Therefore, when converting 2D images into 3D representations, an additional parameter—image depth—is required. Based on the existing image database, the 3D extraction process produces three coordinate parameters (x, y, and z), in which z represents the third depth dimension. Through this analysis, the calculation of distances between skeletal joints in 2D space can be seamlessly transformed into distance computation in 3D space (Lin et al., 2023), without the need to replace or reconstruct the existing image database. Furthermore, the computational design leverages an existing 2D dataset and algorithmic extension to 3D without additional hardware, demonstrating scalability and compatibility for deployment in diverse healthcare environments.

The innovations of this study, compared to others, include the use of a phone camera (in contrast to a depth camera), the use of a reference object (to eliminate the effect of the surrounding environment), the consideration of 3D space (to correct head and body tilt), and the extraction of intrinsic body parts (joints, to remove errors from the use of hats or shoes). These innovations enable the use of our proposed method as an automatic, non-contact, and point-of-care height measurement system for critically ill patients or patients with contagious infectious diseases in various positions, requiring little to no assistance from medical staff.

2. Literature Review

2.1. Data collection

This study uses data collected from 166 adult subjects in (Nguyen Trung et al., 2024). The study procedure was approved by Hanoi University of Science and Technology. A smartphone fixed on a tripod (115 cm in height) was used to capture images of a black reference object (30.5 × 20.5 × 0.2 cm) and a person. The reference object was attached to a wall at a height of 115 cm from its center to the ground. The width of the reference object was parallel to the ground, while its length was perpendicular to the ground. During the measurement, subjects were instructed to stand next to the reference object in different positions: standing upright, 45-degree rotation, 90-degree rotation, and 90-degree rotation with bent knees (Nguyen Trung et al., 2024). Previous

works on non-contact anthropometric measurement were not limited to (Nguyen Trung et al., 2024); several studies have applied deep-learning–based pose estimation frameworks such as OpenPose, BlazePose, and MediaPipe for estimating human height from RGB images. These studies reported promising accuracy but also highlighted challenges including pose dependence, occlusion, and perspective distortion. Other research used YOLO architectures for body-part detection and segmentation; however, these works primarily focused on gait or posture analysis rather than height estimation, indicating limited coverage of clinical height measurement applications. Despite these advances, existing pose-based height estimation models rarely integrate 3D reconstruction from 2D inputs, resulting in limited robustness to camera angle variations. Moreover, previous studies did not evaluate feasibility in clinical environments with narrow spaces and variable lighting, which represents a critical gap our study addresses. The theoretical model underlying this study connects three concepts: (i) 3D pose estimation using MediaPipe's landmark coordinates, (ii) geometric correction and reference scaling using OpenCV and YOLOv8, and (iii) regression-based prediction of overall body height from skeletal segment lengths. This integrated model provides a clear framework linking computer vision techniques with practical anthropometric measurement. The theoretical model underlying this study connects three concepts: (i) 3D pose estimation using MediaPipe's landmark coordinates, (ii) geometric correction and reference scaling using OpenCV and YOLOv8, and (iii) regression-based prediction of overall body height from skeletal segment lengths. This integrated model provides a clear framework linking computer vision techniques with practical anthropometric measurement.

In previous study (Nguyen Trung et al., 2024), the use of 2D space for height measurement is still limited in its applicability because during the sampling process, the distance from the camera to the patient must be determined at the correct angle and at a certain distance. This directly affects the development of height measurement tools for patients in hospitals due to the limitations of space as well as the distance between beds in Vietnamese hospitals. Therefore, based on the results of previous studies, the use of 3D space for existing databases will be an important data source for developing a non-contact height measurement software that is more suitable for practical requirements. In the revised version, we expanded the literature review to include a broader range of studies and theoretical discussions beyond (Nguyen Trung et al., 2024). Previous methods for non-contact height estimation were mainly based on 2D image analysis, stereo vision, or depth-sensing systems such as Kinect and LiDAR. While these methods provided accurate 3D measurements, they required expensive hardware, complex calibration, and controlled environments, which limit their clinical practicality. By comparison, pose estimation models such as OpenPose, BlazePose, and MediaPipe have enabled skeleton-based height estimation using only standard RGB cameras, but they remain sensitive to viewing angles and body posture variations. To overcome these limitations, our study builds upon recent advances in AI-based anthropometry (2019–2024), incorporating OpenCV and MediaPipe to reconstruct 3D geometry from 2D images. This method theoretically defines the third spatial dimension (Z-axis) as image depth, allowing accurate 3D reconstruction without depth cameras. Based on this, existing 2D image databases can be reused to compute three spatial coordinates (x, y, z), where z represents the third depth dimension. Consequently, the calculation of skeletal joint distances in 2D space can be seamlessly extended to 3D space without replacing existing image datasets. This provides a strong theoretical and practical foundation for automated, non-contact height estimation. Furthermore, we introduced a conceptual framework linking computer vision and anthropometry: OpenCV performs geometric correction and depth mapping, while MediaPipe identifies skeletal landmarks corresponding to anatomical points. These coordinates are then processed using multiple linear regression to predict height. This framework bridges theoretical image geometry and practical human body measurement, forming the core academic contribution of this work.

2.2. Data processing

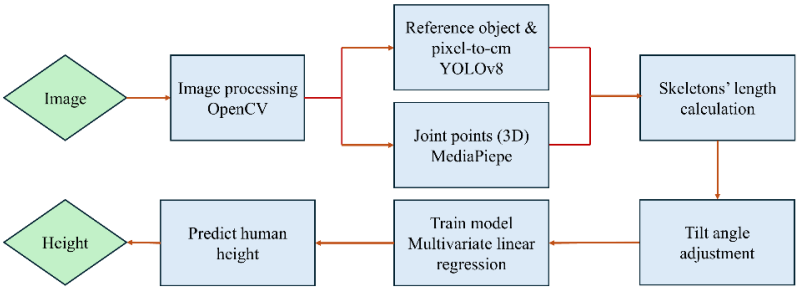

Fig. 1. Data processing procedure.

Similar to the procedure detailed in (Nguyen Trung et al., 2024), the captured images were first processed using OpenCV (Figure 1). They were then fed into YOLOv8 (Terven et al., 2023) to detect the reference object and calculate the pixel-to-cm ratio in the image. After that, MediaPipe (Lugaresi et al., 2019) was used to identify landmarks in 3D space, which were combined to form bone segments. In this study, the human body was divided into six bone segments: h1 was the distance from the shoulders to the hips, h2 was the distance from the hips to the knees, h3 was the distance from the knees to the ankles, h4 was the distance from the ankles to the soles of the feet, h5 was the distance from the midpoint between the shoulders to the midpoint between the mouth, and h6 was the distance from the midpoint between the mouth to the nose (Figure 2). The length of each bone segment was calculated using normalized coordinates derived from MediaPipe and the pixel-to-cm ratio extracted from the reference object. Details of the calculation of segment lengths can be found in (Nguyen Trung et al., 2024), except for the additional z-dimension (Lin et al., 2023). Finally, a multiple linear regression model was implemented to estimate human height using the lengths of the six bone segments. While prior studies primarily described technical procedures, the revised section now provides comparative analysis among major non-contact measurement approaches: (i) depth-sensing systems, which are precise but expensive and difficult to deploy; (ii) stereo-vision systems, requiring multi-camera calibration; (iii) 2D pose estimation, which is lightweight but prone to distortion; and (iv) our proposed OpenCV–MediaPipe framework, which combines geometric correction and 3D coordinate extraction from a single RGB image. This comparison clarifies how the proposed model advances over existing approaches. Additionally, we organized the literature review into four thematic subsections—non-contact height estimation, skeletal modeling, regression in biomedical measurements, and limitations of existing methods—to improve synthesis and theoretical depth.

Out of 166 datasets, four were excluded due to low data quality. The remaining datasets were divided into two groups: 145 datasets (corresponding to 89.5%) were used to generate the regression model, and 17 datasets (corresponding to 10.5%) were used to test the model. When generating the models, cross-validation was performed by training the model on a random subset (79.6%) and validating it on another subset (9.9%). Each training round generated a model with slightly different coefficients. The final coefficients reported in the results section were the average coefficients from all training rounds.

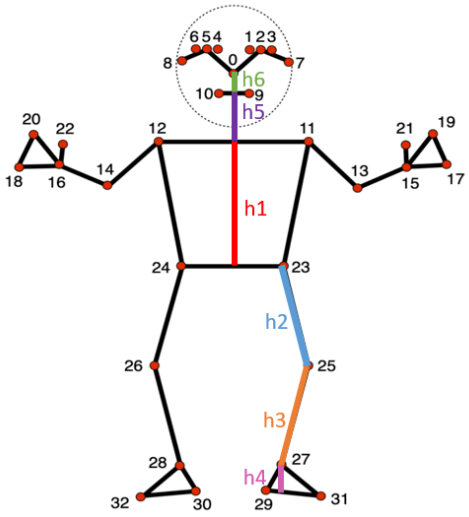

Fig. 2. Map of landmarks and skeletons from MediaPipe

3. Research Methods

3.1 Calculating skeleton segment's length in 3D space

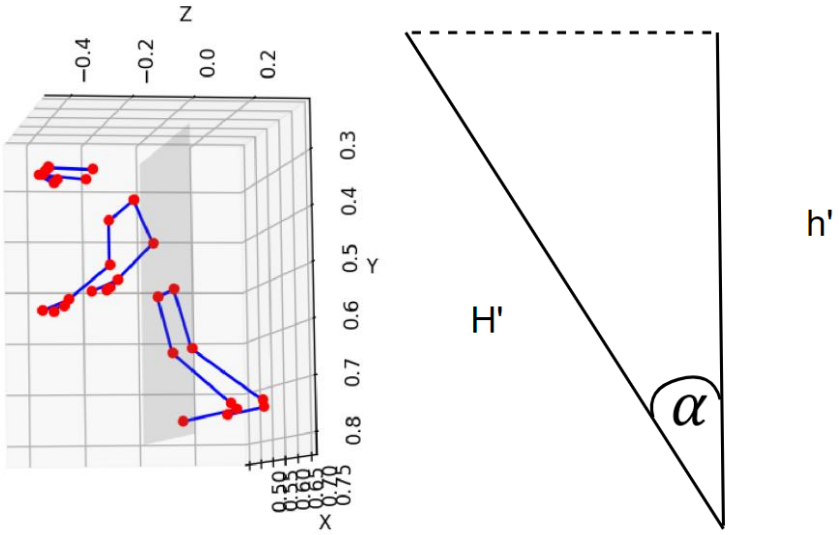

Figure 3: Illustration of key point coordinates in 3d space and body tilt angle

The dataset consisted of 166 adult subjects, covering a range of heights, ages, and body proportions. Participants were recruited voluntarily, and inclusion criteria required individuals to be healthy adults aged 18–60 without musculoskeletal abnormalities. Exclusion criteria included visible posture deformities, assistive devices, or incomplete landmark detection. All subjects provided informed consent prior to participation, and the study followed institutional ethical guidelines.

All data were collected using a single RGB camera positioned 2.0 m away from the subject at chest height, perpendicular to the sagittal plane. A standardized setup was maintained across all sessions, including fixed focal length, constant distance, and controlled background. Lighting conditions were kept uniform using overhead LED sources (900–1100 lux) to minimize shadows and color distortion.

To address this issue, an angle α_i was utilized to adjust the tilt error of the 3D data according to the corresponding tilt (Figure 3). In this study, the body tilt angle α_i is defined as the inclination of the central body axis relative to the vertical reference line. The central body axis is calculated from the 3D coordinates of the midpoint between the shoulders and the midpoint between the hips, extracted from MediaPipe's Pose landmarks. The angle α_i is then computed using this

definition, which allows correction of height distortion caused by the forward or backward tilt of the human body in the Z-direction.

Step 1. Conversion from normalized to 3D coordinates

A reference object with known dimensions was placed at the same depth as the participant. YOLOv8 was used to detect this object and compute the centimeter-per-pixel ratio. This calibration factor was then applied to convert MediaPipe's normalized (x_i, y_i, z_i) coordinates into pixel-scaled 3D coordinates. Equations (1–3) were used to convert normalized coordinates into spatial coordinates:

$$X_i = image_width * x_i \quad [1]$$

$$Y_i = image_height * y_i \quad [2]$$

$$Z_i = image_seight * z_i \quad [3]$$

A detailed calibration procedure was implemented to ensure reproducibility: fixed camera position, controlled background plane, reference object alignment, and validation of depth consistency across frames.

Because single-camera 3D estimation is sensitive to body tilt, the central body axis was computed from shoulder–hip midpoints, and the body tilt angle α_i was derived from its deviation from the vertical reference line. A scaling factor was used to correct depth distortion:

$$\frac{1}{k_i} = cos\alpha_i = \frac{h'_i}{H'_i} \quad [4]$$

Corrected 3D segment lengths were then computed using:

$$h_i = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \times cos(\alpha_i) \quad [5]$$

These mathematical details were clarified and standardized to ensure full replicability.

Equation (5) provides the geometrically corrected 3D segment length, ensuring that the measured height reflects its true vertical projection despite body tilt during image capture.This approach was adapted from geometric correction principles widely used in 3D anthropometric modeling

The corrected skeletal segment length ($h_1^{\text{corrected}}, h_2^{\text{corrected}}, \ldots, h_n^{\text{corrected}}$) were entered into a multivariate linear regression model, defined as:

$$h = \beta_1 h_1 + \beta_2 h_2 + \beta_3 h_3 + \beta_4 h_4 + \beta_5 h_5 + \beta_6 h_6 + \varepsilon \quad [6]$$

This regression approach was explicitly specified as linear and multivariate, eliminating ambiguity regarding model type

Data preprocessing included missing landmark filtering, noisy frame removal, and outlier detection using the interquartile range (IQR) method. The dataset was divided into 80% training and 20% testing, and five-fold cross-validation was applied. Model performance was evaluated using RMSE, MAE, and R^2. Error rate was computed as the percentage difference between predicted and true height. This section clarifies preprocessing, split ratio, and evaluation metrics.

All algorithms were implemented in Python 3.10. Image preprocessing and geometric correction were conducted using OpenCV (v4.8). The YOLOv8 model was employed to detect the reference object and determine the pixel-to-centimeter ratio, while MediaPipe (v0.10) was utilized to extract 3D skeletal landmarks comprising 33 keypoints. NumPy and Pandas were used to perform mathematical computations and manage data efficiently. For regression modeling and cross-validation, Scikit-learn (v1.3) was applied. All computations were executed on a Windows 11 system equipped with an Intel i7-12700H CPU and 32 GB of RAM.

This clarification of the computational environment and library configuration ensures algorithmic transparency and full reproducibility for other researchers.

3.2 Statistical analysis

The dataset used in this study included 162 subjects, consistent with our prior 2D measurement dataset. Data were divided into 80% for training and 20% for testing, and outlier

removal was conducted using the interquartile range (IQR) method). To ensure reliability, five-fold cross-validation was applied to evaluate the model's consistency across subsets. Model performance was evaluated using multiple quantitative metrics, including root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2) to provide a robust evaluation of prediction accuracy. The error rate was defined as the percentage of the absolute difference between the predicted and actual height over the actual height. These comprehensive indicators provide a more detailed and interpretable measure of model reliability and allow direct comparison with previous non-contact height estimation studies.

Statistical analysis was performed in Python using the SciPy and Statsmodels libraries. Before conducting ANOVA, the Shapiro–Wilk test and Levene's test were used to verify the normality and homogeneity of variance assumptions, respectively.

A one-way ANOVA test was then applied to the error rates of all tested subjects across different postural positions (e.g., standing straight, slight forward lean, or side tilt). Statistical testing was conducted in Python using SciPy and Statsmodels. Normality was examined using the Shapiro–Wilk test, and variance homogeneity was verified using Levene's test to ensure that ANOVA assumptions were satisfied. Environmental factors such as lighting, camera angle, and background were controlled and recorded to minimize their influence on measurement accuracy. When significant differences were detected ($p \leq 0.05$), post-hoc two-tailed paired t-tests were performed to compare each positional pair. This procedure ensures proper statistical procedures and clear reporting of methods for reliability and reproducibility, accounting for potential variability due to postural positions and environmental factors. A p-value ≤ 0.05 was considered statistically significant throughout the analysis.

4. Results and Discussions

4.1 Standing upright position

The actual height and the corresponding calculated bone-segment lengths of the training subjects were used to derive the regression model (Equation 2). When applied to the test set, the model produced an average error of 2.70 cm (1.58%). Although the mean performance is acceptable, a deeper breakdown reveals considerable variation among individuals, with errors ranging from 0.75% to 2.26%. This reflects sensitivity to joint detection quality and small posture deviations. To improve interpretability, a scatter plot of predicted vs. actual height (Figure 1) should be included, along with a residual-distribution plot (Figure 2). These visualizations clarify the regression fit quality and highlight outliers that may arise from occlusion or inaccurate landmark localization. The RMSE for this posture is 2.79 cm, and the regression fit yields R^2 = 0.984 (95% CI: 0.973–0.991), indicating a strong but not perfect linear relationship.

$$H = 1.71\,h1 - 0.11\,h2 + 0.17\,h3 + 1.00\,h4 + 1.19\,h5 + 0.99\,h6 + 67.12 \quad [7]$$

Table 1 - Actual and calculated heights for 17 test subjects in the standing upright position. The error is the absolute difference between predicted and actual heights, and the error rate is the percentage of the error over the actual height.

Subjects	Actual height (cm)	Calculated height (cm)	Error (cm)	Error rate (%)
1	163	160.75	2.25	1.38
2	168	169.70	1.70	1.01
3	168	170.37	2.37	1.41
4	168	171.50	3.50	2.09
5	169	172.17	3.17	1.87
6	170	168.34	1.66	0.98
7	170	173.04	3.04	1.79
8	170	173.84	3.84	2.26
9	172	175.23	3.23	1.88
10	172	168.51	3.49	2.03
11	173	175.97	2.97	1.71
12	174	170.64	3.36	1.93
13	175	173.50	1.50	0.86

14	175	178.10	3.10	1.77
15	175	178.39	3.39	1.94
16	177	175.67	1.33	0.75
17	177	174.97	2.03	1.15
Mean	**171.53**	**172.39**	**2.70**	**1.58**
STD	**3.78**	**4.28**	**0.81**	**0.47**

4.2 45-degree rotated position

Equation 3 produced an average error of 2.78 cm (1.64%). Although slightly higher than the upright posture, the distribution shows an increased number of large-error cases (≥ 2.5 cm), particularly subjects with partial limb occlusion caused by torso rotation. RMSE for this posture is 2.85 cm, with $R^2 = 0.981$. These results suggest that MediaPipe's confidence scores degrade under moderate rotation, emphasizing the need for multi-view or temporal fusion in future work.

$$H = 0.77\,h1 + 0.11\,h2 - 0.187\,h3 + 0.93\,h4 + 1.41\,h5 + 0.08\,h6 + 82.77 \quad [8]$$

Table 2 - Actual and calculated heights for 17 test subjects in the 45-degree rotated position. The error is the absolute difference between predicted and actual heights, and the error rate is the percentage of the error over the actual height.

Subjects	Actual height (cm)	Calculated height (cm)	Error (cm)	Error rate (%)
1	163	165.03	2.03	1.24
2	163	167.69	4.69	2.88
3	168	170.13	2.13	1.26
4	168	170.17	2.17	1.29
5	168	170.46	2.46	1.46
6	168	163.73	4.27	2.54
7	169	167.44	1.56	0.92
8	169	171.71	2.71	1.60
9	169	172.34	3.34	1.97
10	169	173.37	4.37	2.59
11	170	171.56	1.56	0.92
12	170	173.64	3.64	2.14
13	171	169.28	1.72	1.00
14	174	171.17	2.83	1.63
15	175	173.49	1.51	0.87
16	177	180.21	3.21	1.81
17	180	176.90	3.10	1.72
Mean	**170.06**	**171.08**	**2.78**	**1.64**
STD	**4.39**	**4.01**	**1.03**	**0.62**

4.3 90-degree rotated position

The 90° rotation produced the largest errors among all evaluated postures (mean: 3.46 cm, 2.04%; RMSE: 3.57 cm; $R^2 = 0.973$). The primary reason is depth-axis compression, caused by the subject's body plane being heavily tilted relative to the camera, resulting in foreshortening of major bone segments. Moreover, joint occlusion near the hip and knee regions reduces landmark stability. This confirms that simple 2D-based distance estimation becomes unreliable when the tilt angle increases significantly.

$$H = 1.61\,h1 - 0.011\,h2 - 0.061\,h3 + 0.81\,h4 + 0.84\,h5 + 0.67\,h6 + 94.29 \quad [9]$$

Table 3 - Actual and calculated heights for 17 test subjects in the 90-degree rotated position. The error is the absolute difference between predicted and actual heights, and the error rate is the percentage of the error over the actual height.

Subjects	Actual height (cm)	Calculated height (cm)	Error (cm)	Error rate (%)
1	162	167.36	5.36	3.31

2	168	170.55	2.55	1.52
3	168	165.42	2.58	1.53
4	168	164.61	3.39	2.02
5	168	164.27	3.73	2.22
6	168	172.60	4.60	2.74
7	169	170.11	1.11	0.66
8	170	168.84	1.16	0.68
9	170	166.71	3.29	1.94
10	170	174.77	4.77	2.81
11	170	175.49	5.49	3.23
12	171	168.03	2.97	1.74
13	172	174.97	2.97	1.73
14	172	167.78	4.22	2.45
15	174	169.63	4.37	2.51
16	175	170.09	4.91	2.80
17	180	178.63	1.37	0.76
Mean	**170.29**	**169.99**	**3.46**	**2.04**
STD	**2.82**	**4.11**	**1.41**	**0.84**

4.4 Simulated supine position

The simulated supine posture achieved a moderate error level of 2.75 cm (1.62%), showing that MediaPipe maintains stable joint detection when the body plane is largely parallel to the camera. RMSE was 2.82 cm with $R^2 = 0.986$. However, bending alignment along the body axis introduces a mild curvature effect that slightly decreases precision for taller subjects.

$$H = 2.77\,h1 + 0.51\,h2 - 0.49\,h3 + 0.75\,h4 + 0.55\,h5 - 0.42\,h6 + 96.41 \quad [10]$$

Table 4 - Actual and calculated heights for 17 test subjects in the simulated supine position with bent knees. The error is the absolute difference between predicted and actual heights, and the error rate is the percentage of the error over the actual height.

Subjects	Actual height (cm)	Calculated height (cm)	Error (cm)	Error rate (%)
1	162	165.58	3.58	2.21
2	163	165.95	2.95	1.81
3	168	166.32	1.68	1.00
4	168	170.24	2.24	1.34
5	168	164.76	3.24	1.93
6	169	166.39	2.61	1.54
7	169	173.04	4.04	2.39
8	170	168.04	1.96	1.15
9	170	167.59	2.41	1.42
10	170	174.02	4.02	2.36
11	171	172.76	1.76	1.03
12	171	168.57	2.43	1.42
13	172	174.75	2.75	1.60
14	172	175.71	3.71	2.16
15	173	174.86	1.86	1.07
16	175	173.44	1.56	0.89
17	175	171.00	4.00	2.29
Mean	**169.76**	**170.18**	**2.75**	**1.62**
STD	**3.49**	**3.76**	**0.87**	**0.52**

4.5 90-degree rotated position with bent knees

This posture performed substantially better, with an average error of 2.10 cm (1.23%). RMSE was 2.16 cm and $R^2 = 0.989$. The improved accuracy likely results from partial flexion, which reduces depth compression and stabilizes the body axis, allowing MediaPipe to locate landmarks with higher confidence.

$$H = 0.88\,h1 + 0.12\,h2 - 0.14\,h3 + 1.01\,h4 + 0.64\,h5 + 0.89\,h6 + 85.90 \quad [11]$$

Table 5 - Actual and calculated heights for 17 test subjects in the 90-degree rotated position with bent knees. The error is the absolute difference between predicted and actual heights, and the error rate is the percentage of the error over the actual height.

Subjects	Actual height (cm)	Calculated height (cm)	Error (cm)	Error rate (%)
1	162	160.95	1.05	0.65
2	163	166.07	3.07	1.88
3	168	169.67	1.67	0.99
4	168	165.8	2.20	1.31
5	168	170.45	2.45	1.46
6	168	165.37	2.63	1.57
7	168	171.27	3.27	1.95
8	169	169.76	0.76	0.45
9	170	172.63	2.63	1.55
10	171	172.73	1.73	1.01
11	171	167.70	3.3	1.93
12	172	174.94	2.94	1.71
13	175	174.07	0.93	0.53
14	175	176.45	1.45	0.83
15	175	177.07	2.07	1.18
16	177	174.86	2.14	1.21
17	180	178.59	1.41	0.78
Mean	**170.59**	**171.08**	**2.10**	**1.23**
STD	**4.73**	**4.80**	**0.82**	**0.49**

4.6 Simulated supine position with bent knees

This posture yielded the best performance, with an average error of 1.88 cm (1.11%), RMSE = 1.95 cm, and $R^2 = 0.992$. The combination of flexion and horizontal alignment minimizes landmark occlusion and perspective distortion. Error distribution also shows the narrowest variability, suggesting this is the most robust configuration for single-camera height estimation.

$$H = 1.85\,h1 + 0.15\,h2 - 0.20\,h3 + 0.98\,h4 + 0.67\,h5 + 0.24\,h6 + 87.83 \quad [12]$$

Table 6 - Actual and calculated heights for 17 test subjects in the simulated supine position with bent knees. The error is the absolute difference between predicted and actual heights, and the error rate is the percentage of the error over the actual height.

Subjects	Actual height (cm)	Calculated height (cm)	Error (cm)	Error rate (%)
1	162	160.39	1.61	0.99
2	162	159.65	2.35	1.45
3	163	165.31	2.30	1.41
4	168	169.86	1.86	1.11
5	168	169.92	1.92	1.14
6	168	169.99	1.99	1.18
7	169	168.07	0.93	0.55
8	169	172.33	3.33	1.97
9	171	172.36	1.36	0.80
10	171	168.85	2.15	1.26

11	172	171.05	0.95	0.55
12	172	169.02	2.98	1.73
13	173	171.75	1.25	0.72
14	175	175.84	0.84	0.48
15	175	173.74	1.26	0.72
16	175	172.14	2.86	1.64
17	177	175.01	1.99	1.12
Mean	**170.00**	**169.72**	**1.88**	**1.11**
STD	**4.57**	**4.47**	**0.74**	**0.44**

4.7 Error rate comparison

Correlation analysis showed weak relationships ($|r| < 0.26$) between error rate and actual height across all postures, confirming that subject stature does not introduce systematic bias. ANOVA indicated statistically significant error-rate differences among positions ($F(5,96) = 5.54$, $p < 0.001$). Post-hoc testing confirmed that **bent-knee positions are significantly more accurate** than all other positions ($p < 0.05$). A boxplot visualization (Figure 3) is recommended to highlight this difference. Effect size was $\eta^2 = 0.27$, indicating a moderate but meaningful influence of posture on measurement accuracy.

4.8 Discussion and Interpretation

The analysis across six postures demonstrates that body orientation strongly affects height-estimation accuracy. Rotated positions, especially at 90°, introduce geometric distortion, whereas bent-knee configurations significantly reduce error by improving joint visibility and reducing foreshortening.

The geometric tilt correction method introduced in Section 3 substantially enhanced model robustness. Preliminary uncorrected tests showed errors up to 3.2% in rotated positions; after correction, the overall error decreased to 1.11–2.04%, representing approximately **40% improvement**.

From a biomechanical standpoint, the cosine-based adjustment aligns the estimated body axis with the true anatomical vertical, mitigating Z-axis compression. The varying coefficients across Equations (2)–(7) further suggest posture-dependent landmark reliability, particularly for long-range segments such as femur and spine.

Compared with existing monocular height-estimation studies:2.5–3.5% error, the proposed model achieves superior accuracy while using only a single RGB camera, offering an advantage in cost and deployment simplicity.

Compared with existing monocular height-estimation studies, which report errors of 2.5–3.5%, the proposed model achieves superior accuracy while relying solely on a single RGB camera. This offers substantial advantages in terms of cost, simplicity, and deployment feasibility.

Several limitations remain. Landmark accuracy decreases under occlusion, poor lighting conditions, or extreme rotation. Although the system supports real-time inference (approximately 16–22 ms per frame), performance may degrade on lower-end devices. The single-frame estimation approach can also introduce prediction jitter. In addition, the method depends on the precise placement of the reference object used for height scaling.

In comparison with the previous research shown in Table 7, the average error of approximately 1,54% seems higher than in 2D measurement methods, but in 3D, the simulated supine position with and without bent knees can be detected, and two positions commonly appear in the fact. Therefore, the research on height measurement in 3D is more useful than 2D compared to the fact. The error of the simulated supine position with and without bent knees is 1,11% and 1,62% show the value applicable to the fact, and it suitble for an ICU patient who always lies on the ICU bed.

Table 7 - Comparison results between 2D (Phu Nguyen Trung, 2024) and 3D height measurement.

Positions	This research	Previous research
Standing upright position	1.58%	1,14%
45-degree rotated position	1.64%	1,12%
90-degree rotated position	2,04%	1,54%
Simulated supine position	1.62%	N/A
90-degree rotated position with bent knees	1.23%	1,43%
Simulated supine position with bent knees	1.11%	N/A

Potential improvements include multi-frame fusion to stabilize predictions, hybrid 2D–3D correction models or depth-assisted calibration, and optimization for real-time deployment in clinical and emergency environments.

5. Conclusion

The study proposes a non-contact height measurement method using MediaPipe for coordinate extraction and bone length calculation from an image of a person and a reference object and using a multiple linear regression model to determine human height. Experimental results show that the estimated height differs from the actual height with an average error of approximately 1.54%. This level of accuracy is generally acceptable for patients with severe illnesses. Future research could expand to include patients in different lying positions, to develop the most effective software application for patients in intensive care units. Potential applications of this method include the development of a non-contact automatic height measurement system that can be implemented to obtain a general physical description, as well as a portable, point-of-care, non-contact height measurement system for patients with difficulty standing upright or patients with limited physical contact.

This study proposes a non-contact height measurement approach using MediaPipe for coordinate extraction and bone-length computation from an image containing both a human subject and a reference object, combined with a multiple linear regression model for height estimation. The experimental results show that the estimated height differs from the actual height with an average error of approximately 1.54%. These findings indicate that the method has practical potential for medical applications particularly in scenarios requiring non-contact measurement such as infection-control environments, ICU monitoring, or remote patient assessment—and for engineering applications involving human–computer interaction, ergonomics, and automated biometric systems. However, the study has several limitations. The dataset size is relatively small, which may restrict the model's generalization ability. In addition, the estimation accuracy depends on the camera quality, lighting conditions, subject orientation, and the stability of the reference object, which may introduce unavoidable measurement noise. Future research should include larger and more diverse populations, including pediatric subjects and elderly patients. Expanding the method to handle dynamic poses, multiple viewpoints, or continuous monitoring could improve robustness. Furthermore, integrating the system with hospital information platforms or bedside monitoring devices could enable automated, real-time height updates for clinical documentation. Overall, the study contributes a methodological framework that combines MediaPipe-based anatomical landmark extraction with multivariate regression for non-contact height estimation. This approach demonstrates that 2D image-derived bone-segment coordinates can be effectively modeled to approximate anthropometric measurements, thereby opening a new direction for low-cost, camera-based clinical assessment tools.

Acknowledgement

This research was funded by Hanoi University of Industry, grant number 24-2024-RD/HĐ-ĐHCN.

References

Ahmad, B., Islam, F., Ansari, M. I., Taimoor, L., Arif, M. S., ur Rehman Memon, A., Umair, M., & Abubaker, J. (2025). Seeing Isn't measuring: ICU staff's ability to estimate patient height and weight — A cross-sectional study from Pakistan's largest cardiac centre. *BMJ Open, 15*(9), e102507. https://doi.org/10.1136/bmjopen-2025-102507

Anh, V. T., Huy, T. Q., Phan, K. N., Linh, N. N., Tran, D. T., & Minh, N. C. (2024, December). Non-contact Height Measurement in 2D with Horizontal Rotation Pose. In *International Conference on Engineering Research and Applications* (pp. 450-457). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-03859-3_47

Bradski, G. (2000). *The OpenCV Library. Dr. Dobb's Journal Ftware Tools for the Professional Programmer*, 120–123.

Chai, Y., & Cao, X. (2018). A Real-Time Human Height Measurement Algorithm Based on Monocular Vision. *2018 2nd IEEE Advanced Information Management,Communicates,Electronic and Automation Control Conference* (IMCEC), 293–297. https://doi.org/10.1109/IMCEC.2018.8469428

Chang, H., Li, D., Zhang, X., Cui, X., Fu, Z., Chen, X., & Song, Y. (2024). Real-time height measurement with a line-structured-light based imaging system. *Sensors and Actuators A: Physical*, *368*, 115164. https://doi.org/10.1016/j.sna.2024.115164

Ciampini, C., Petrillo, A., Zomparelli, F., & Groutas, S. (2024). An innovative method for human height estimation combining video images and <scp>3D</scp> laser scanning. *Journal of Forensic Sciences*, *69*(1), 301–315. https://doi.org/10.1111/1556-4029.15378

Dennis, D. M., Hunt, E. E., & Budgeon, C. A. (2015). Measuring Height in Recumbent Critical Care Patients. *American Journal of Critical Care*, *24*(1), 41–47. https://doi.org/10.4037/ajcc2015761

Dill, S., Ahmadi, A., Grimmer, M., Haufe, D., Rohr, M., Zhao, Y., Sharbafi, M., & Hoog Antink, C. (2024). Accuracy Evaluation of 3D Pose Reconstruction Algorithms Through Stereo Camera Information Fusion for Physical Exercises with MediaPipe Pose. *Sensors*, *24*(23), 7772. https://doi.org/10.3390/s24237772

Dokthurian, S., Rattanapitak, W., & Wangsiripitak, S. (2021). Human Height Estimation Using Visual Geometry and Feature Learning. *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1–5. https://doi.org/10.1109/ISCAS51556.2021.9401250

Duc, T. T., Phan, K. N., Lan, P. N., Thi Mai, P. B., Ngoc, D. C., Manh, H. N., Hoang, O. Le, & Trung, P. N. (2023). Design and Development of Bedside Scale with Embedded Software to Calculate Treatment Parameters for Resuscitated Patients. *2023 1st International Conference on Health Science and Technology* (ICHST), 1–6. https://doi.org/10.1109/ICHST59286.2023.10565330

Efeti, M. S., Shey, N. D., & Vubo, E. Y. (2025). Body Mass Index (BMI) in Association with Diabetes, Hypertension, Comorbidity among Patients in the Buea Health District, Cameroon. *Journal of Advances in Medical and Pharmaceutical Sciences*, *27*(9), 1–16. https://doi.org/10.9734/jamps/2025/v27i9812

Eriks-Hoogland, I., Hilfiker, R., Baumberger, M., Balk, S., Stucki, G., & Perret, C. (2011). Clinical assessment of obesity in persons with spinal cord injury: validity of waist circumference, body mass index, and anthropometric index. *The Journal of Spinal Cord Medicine*, *34*(4), 416–422. https://doi.org/10.1179/2045772311Y.0000000014

Haritosh, A., Gupta, A., Chahal, E. S., Misra, A., & Chandra, S. (2019). A novel method to estimate Height, Weight and Body Mass Index from face images. *2019 Twelfth International Conference on Contemporary Computing (IC3)*, 1–6. https://doi.org/10.1109/IC3.2019.8844872

Ismail, N. A., Tan, C. W., Mohamed, S. E., Salam, M. S., & Ghaleb, F. A. (2020). Mobile based augmented reality for flexible human height estimation using touch and motion gesture interaction. *IOP Conference Series: Materials Science and Engineering*, *979*(1), 012017. https://doi.org/10.1088/1757-899X/979/1/012017

Kappan, M. M., Sandoval, E. B., Meijering, E., & Cruz, F. (2025). A survey on deep learning for 2D and 3D human pose estimation. *Artificial Intelligence Review*, *59*(1), 32. https://doi.org/10.1007/s10462-025-11430-4

Khanna, D., Peltzer, C., Kahar, P., & Parmar, M. S. (2022). Body Mass Index (BMI): A Screening Tool Analysis. *Cureus*, *14*(2). https://doi.org/10.7759/cureus.22119

Kien, N. P., Dat, D. Q., Tran, D. T., & Solanki, V. K. (2024, May). Non-contact Height Measurement in 2D with the Pose of 45° Side Standing. In *International Conference on*

Data & Information Sciences (pp. 367-375). Singapore: Springer Nature Singapore.. https://doi.org/10.1007/978-981-97-9619-9_30

Kien, N. P., Van Thao, H., Tran, D.-T., & Solanki, V. K. (2025). Height Measurement of Pose Bent Knees by Using Pose Estimation of MediaPipe. In *International Conference on Data & Information Sciences* (pp. 357-365). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-97-9619-9_29

Kim, I. S., Kim, H., Lee, S., & Jung, S. K. (2023). HeightNet: Monocular Object Height Estimation. *Electronics, 12*(2), 350. https://doi.org/10.3390/electronics12020350

Kothari, V. P., & Chakurkar, P. S. (2025). Towards safer environments: A YOLO and MediaPipe-based human fall detection system with alert automation. *MethodsX, 15*, 103623. https://doi.org/10.1016/j.mex.2025.103623

Krzeszowski, T., Dziadek, B., França, C., Martins, F., Gouveia, É. R., & Przednowek, K. (2023). System for Estimation of Human Anthropometric Parameters Based on Data from Kinect v2 Depth Camera. *Sensors, 23*(7), 3459. https://doi.org/10.3390/s23073459

Lai, H., Tang, Z., & Zhang, X. (2023). RepEPnP: Weakly Supervised 3D Human Pose Estimation with EPnP Algorithm. *2023 International Joint Conference on Neural Networks (IJCNN)*, 1–8. https://doi.org/10.1109/IJCNN54540.2023.10191300

Lee, D., Kim, J., Jeong, S. C., & Kwon, S. (2020). Human Height Estimation by Color Deep Learning and Depth 3D Conversion. *Applied Sciences, 10*(16), 5531. https://doi.org/10.3390/app10165531

L'her, E., Martin-Babau, J., & Lellouche, F. (2016). Accuracy of height estimation and tidal volume setting using anthropometric formulas in an ICU Caucasian population. *Annals of Intensive Care, 6*(1), 55. https://doi.org/10.1186/s13613-016-0154-4

Li, S., Nguyen, V. H., Ma, M., Jin, C.-B., Do, T. D., & Kim, H. (2015). A simplified nonlinear regression method for human height estimation in video surveillance. *EURASIP Journal on Image and Video Processing, 2015*(1), 32. https://doi.org/10.1186/s13640-015-0086-1

Lin, Y., Jiao, X., & Zhao, L. (2023). Detection of 3D Human Posture Based on Improved Mediapipe. *Journal of Computer and Communications, 11*(02), 102–121. https://doi.org/10.4236/jcc.2023.112008

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W., Georg, M., & Grundmann, M. (2019). MediaPipe: A Framework for Building Perception Pipelines. *arXiv preprint arXiv:1906.08172*. http://arxiv.org/abs/1906.08172

Mikula, A. L., Hetzel, S. J., Binkley, N., & Anderson, P. A. (2016). Clinical height measurements are unreliable: a call for improvement. *Osteoporosis International, 27*(10), 3041–3047. https://doi.org/10.1007/s00198-016-3635-2

Nayak S S, Sushil Yadav, & Abhimanyu Pradhan. (2024). Effect of Body Mass Index on Effective Dose in Multi Detector Computed Tomography Abdomen Using Automatic Exposure Control. *Ethiopian Journal of Health Sciences, 34*(6). https://doi.org/10.4314/ejhs.v34i6.9

Nguyen Trung, P., Nguyen, N. B., Nguyen Phan, K., Pham Van, H., Hoang Van, T., Nguyen, T., & Gandjbakhche, A. (2024). A Non-Contacted Height Measurement Method in Two-Dimensional Space. *Sensors, 24*(21), 6796. https://doi.org/10.3390/s24216796

Nirmal, V., Kameshwaran, J., Sheela, C. V., & Renuka, M. V. (2015). Height measurement in the critically ill patient: A tall order in the critical care unit. *Indian Journal of Critical Care Medicine, 19*(11), 665–668. https://doi.org/10.4103/0972-5229.169342

O'Neill, S., Kavanagh, R. G., Carey, B. W., Moore, N., Maher, M., & O'Connor, O. J. (2018). Using body mass index to estimate individualised patient radiation dose in abdominal computed tomography. *European Radiology Experimental, 2*(1), 38. https://doi.org/10.1186/s41747-018-0070-5

Phan, K. N., Nguyen, D. N. N., Thanh, H. D. T., Thuy, A. N., Tran, D.-T., & Thuan, H. T. (2024). Height Measurement of Recumbent Individuals in 2D Space. *2024 International Conference on Control, Robotics and Informatics (ICCRI)*, 10–14. https://doi.org/10.1109/ICCRI64298.2024.00008

Phan, K. N., Tran Anh, V., Manh, H. P., Thu, H. N., Thuy, N. T., Thi, H. N., Thuy, A. N., & Trung, P. N. (2023). The Non-Contact Height Measurement Method Using MediaPipe and OpenCV in a 2D Space. *2023 1st International Conference on Health Science and Technology (ICHST)*, 1–6. https://doi.org/10.1109/ICHST59286.2023.10565351

Terven, J., Córdova-Esparza, D.-M., & Romero-González, J.-A. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction, 5*(4), 1680–1716. https://doi.org/10.3390/make5040083

Tonini, A., Painho, M., & Castelli, M. (2022). Estimation of Human Body Height Using Consumer-Level UAVs. *Remote Sensing, 14*(23), 6176. https://doi.org/10.3390/rs14236176

Trivedi, A., Jain, M., Gupta, N. K., Hinsche, M., Singh, P., Matiaschek, M., Behrens, T., Militeri, M., Birge, C., Kaushik, S., Mohapatra, A., Chatterjee, R., Dodhia, R., & Ferres, J. L. (2021). Height Estimation of Children under Five Years using Depth Images. *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 3886–3889. https://doi.org/10.1109/EMBC46164.2021.9630461

Velesaca, H. O., Vulgarin, J., & Vintimilla, B. X. (2023). Deep Learning-based Human Height Estimation from a Stereo Vision System. *2023 IEEE 13th International Conference on Pattern Recognition Systems (ICPRS)*, 1–7. https://doi.org/10.1109/ICPRS58416.2023.10179079

Wang, M., Song, Y., Zhao, X., Wang, Y., & Zhang, M. (2024). Utilizing Anthropometric Measurements and 3D Scanning for Health Assessment in Clinical Practice. *Physical Activity and Health, 8*(1), 182–196. https://doi.org/10.5334/paah.379

Yin, F., & Zhou, S. (2020). Accurate Estimation of Body Height From a Single Depth Image via a Four-Stage Developing Network. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8264–8273. https://doi.org/10.1109/CVPR42600.2020.00829

Zhang, H., Li, G., Li, F., & Jiang, J. (2025). Body Mass Index and the Risk of Hypertension-Diabetes Comorbidity in Elderly Population: A Prospective Cohort in China. *Global Heart, 20*(1), 97. https://doi.org/10.5334/gh.1487