Analysis of the skeleton of human movement for orthopedics tasks

Olga Nedzved
Belarusian State University
Minsk, Belarus
onedzved@bsu.by

Shiping Ye
Zhejiang Shuren University
Hangzhou, China
zjsruysp@163.com
ORCID 0000-0002-9771-7168

Chaoxiang Chen
Zhejiang Shuren University
Hangzhou, China
ccx@zjsru.edu.cn

Viktor Anosov
The Republican Scientific and Practical Centre Center of Traumatology and Orthopaedics,
Minsk, Belarus

Mikhail Gerasimenko
The Republican Scientific and Practical Centre Center of Traumatology and Orthopaedics,
Minsk, Belarus

Oleg Sakalouski
The Republican Scientific and Practical Centre Center of Traumatology and Orthopaedics,
Minsk, Belarus

***Abstract*—last time a few algorithm for human pose detection was developed. As rule, the basic element for pose description is skeleton. It is possible extract many important information from such object for orthopedics task. In this paper algorithm for automatic estimation of walking motion is proposed on base reconstruction human skeleton and definition of harmonic component of walking**

Keywords—walking component, motion estimation, skeleton, automation of orthopedics analysis

I. INTRODUCTION

Walking is the main way of human movement that ensures the fulfillment of many everyday tasks. Pathology of walking is one of the most common disorders of motor function for people of all ages. Gait can disorders be caused by various causes, including diseases of the musculoskeletal system, neurological disorders, injuries, etc. Various methods are used to diagnose and monitor walking pathology, however, assessment of walking pathology remains a complex and subjective task, especially when it comes to a qualitative assessment of movement in real time.

Assessing a person's positioning is a problem of localizing anatomical key points, which in physical terms are the joints points of the human body. This problem is mainly focused on finding anatomical key points, combining them into a complete skeleton, and tracking it throughout the video sequence. Pose recognizing in video is a complicated problem due to occlusions, properties of joints, the unification of individual parts of the body in the limb, and later into the skeleton. The complexity increases with the number of people due to physical contacts between them, which imposes significant restrictions on the mode of execution in real time.

Currently, the analysis of walking pathologies is an important task in the medical industry. It allows you to assess the patient's condition, identify possible disorders in the musculoskeletal system and make recommendations for treatment and rehabilitation. However, the traditional method of analyzing walking using video sequences is time-consuming and requires considerable time to perform calculations.

This article describes an algorithm for automatic processing of video data for the analysis of walking pathologies using computer vision and machine learning methods. This algorithm will improve the quality of diagnostics and increase its speed, as well as provide automation of medical research. The results of the study can be used in clinical practice to improve and accelerate the diagnosis of diseases.

The main purpose of the work is to describe a pipeline for automatic calculation of characteristics for a method for diagnosing walking disorders.

II. METHODS FOR DETERMINING OF GAIT DISORDERS

Human walking is a cycle, the main stages of which are: support on the heel, support on the entire foot, support on the front part and transfer of the leg (Figure 1). Each stage is important for walking analysis, because it determines the step frequency, load distribution, the angle at which the foot is placed, and so on [1].

Many scales and marks are employed for the examination of walking, for example Edinburgh Scale and Edinburgh visual gait score is an evaluation instrument that assesses each component of gait using on-screen sketching and measuring tools, software, and video cameras to deliver 3D video gait analysis [1].

The analysis of the patient's gait disorders includes registration of walking from four sides simultaneously. After video recording of the step cycle with subsequent computer processing of the image, the temporal and spatial characteristics of the step are calculated. Markers in the form of lines are placed along the entire axis of the thighs and lower legs in the lateral, frontal and horizontal planes.

The study of a person's gait allows recognition of movement characteristics, establish typical gait patterns, identify conditions that cause pain, and apply and evaluate therapies to eliminate disorders [5].

Machine vision algorithms have several applications for detecting human motion characteristics, starting with looking for a person in a frame and creating a rectangular area around them in 2- or 3-dimensional space.

The task of finding out the posture of a person when moving is more difficult. This is difficult, because even in the most basic model, the size of the vector that determines the position of a person in three dimensions includes more than 100 parameters [6]. On Figure 1 shows a conditional vector representation of the human skeleton in two dimensions [6].

Fig. 1. An example of representing a human pose as a vector

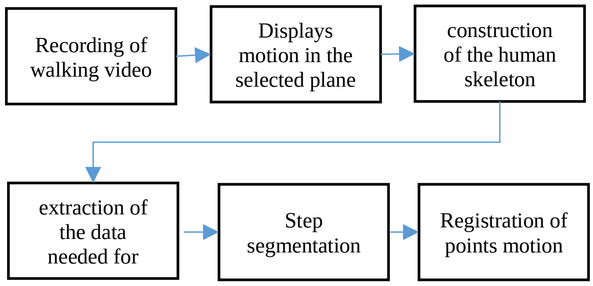

Fig. 2. The diagram of the process of the analysis of gait pathology automatization

To obtain a set of trajectories for points corresponding to the nodes of the skeletal model, it is necessary to sequentially apply algorithms to a series of video stream frames [6]. Since each frame is analyzed separately, the resulting trajectories may contain defects and artifacts of various kinds (outliers, missing trajectory points, mixed up points of adjacent trajectories). Therefore, it is important to develop a model that provides smooth transitions between video slides.

Currently, various algorithms for constructing the human skeleton are used:

OpenPose algorithm [7] is used to identify important parts of the human body, including the shoulders, elbows, wrists, pelvis, and knees. This method uses neural networks. The human skeleton is subsequently built by connecting these key points.

AlphaPose algorithm [7] also uses neural networks to identify key points in the body and join them together to form a skeleton. In addition, it can recognize body movements and gestures.

Kinect Fusion algorithm [8] uses depth cameras to collect information about the movements and shape of the human body. Then, using reconstruction and data processing methods, an accurate three-dimensional human skeleton is created.

SkelNet algorithm [8] also uses depth cameras to record information about the movements and shape of the human body. After that, using data processing methods and machine learning, an accurate three-dimensional human skeleton is created.

BlazePose algorithm [8] uses a large amount of training data to achieve high accuracy based on the ResNet and Hourglass neural network architecture. It has high processing speed and can process both 2D and 3D images.

YOLOv7 Pose algorithm [6] is designed to find and track the key points of the body. It can process real-time video quickly thanks to its YOLO (You Only Look Once) neural network design system. The system uses deep learning and a large amount of training data to improve accuracy.

However, off-the-shelf methods for building a human model need to be improved so that they can calculate all gait indicators without the participation of a doctor.

III. Description Of Algorithms

On Figure 2 a diagram of the process of automatisation the diagnosis of gait disorders is presented.

According to this scheme, a video of the patient's walking in one of the planes is recorded. Algorithm testing was performed for video obtained in the sagittal plane. After receiving a video fragment, a human skeleton is built for each frame using the Mediapipe Pose (Blaze Pose) neural network [7] and the positions of key points are determined. Next, the movement of key points is recorded, and discrete functions of changing the x and y coordinates are constructed for each key point depending on the frame number. The image is then filtered to improve the quality of the analysis. After that, based on the values of discrete functions, segmentation of the cycle corresponding to the step is performed. As a result, doctors receive the necessary information about the position of the limbs at different periods of the step cycle.

To determine the key points, there are a large number of neural networks. They build the human skeleton from a frame/set of frames. The most popular neural networks include YOLOv7 POSE, Mediapipe Pose and Open Pose.

IV. Automation Of The Method Of Diagnostics Of Gait Disorders

In the method of diagnostics gait disorders, it is necessary to quantitatively measure the angular characteristics of the step cycle, as well as the duration of certain periods. Then it is necessary to analyze these characteristics for the presence of pathologies. table 1 shows the data that needs to be automatically calculated for further pathology detection.

TABLE II. Data necessary for the diagnosis of gait disorders (step length 0.8m, duration of the step cycle 1.1s, duration of the support phase 0.75s, duration of the transfer phase 0.35s)

Key joint	**Sagittal plane**					
	Initial contact	***Double support period (complete-tion)***	***Support period***	***Double support period (begin-ning)***	***Push***	***Transfer period (vertical position of the tibia)***
Hip	17	16	-4	-18	-2	30
Knee	22	24	12	16	55	60
Ankle	-5	0	0	5	-30	-15

The method of diagnosing walking disorders uses a treadmill, as well as 4 video cameras that capture the patient's walking process in different planes: sagittal (lateral projection), frontal (rear view), frontal (front view), horizontal (top view). In this paper, the analysis of video obtained from a camera that captures movement only in the sagittal plane is performed (Figure 3).

Fig. 3. Diagram of the registration method for the diagnosis of walking disorders: sagittal plane for step cycle estimation (lateral projection) (1); frontal plane for assessing foot support and angle of movement (rear view) (2); frontal plane for evaluating the frontal characteristics of the step cycle (front view) (3); horizontal plane for evaluating the horizontal characteristics of the step cycle (top view) (4)

In order to automatically receive the data necessary to determine the diagnosis, the following algorithm was proposed:

1. The video fragment is divided into frames using the python OpenCV library.

2. For each frame a skeleton of the patient is built.

3. For each frame x, y coordinates are recorded for the following body parts : left and right heel, left and right toe, left and right knee, left and right hip, angle between heel and treadmill in the sagittal plane, both legs.

4. The obtained discrete functions of changes in the coordinates of the limbs are filtered.

5. The number of the frame at which the step ends is determined by analyzing local extremes.

6. Step segmentation is performed using the analysis of local extremes, as well as the algorithm of dynamic transformation of the timeline and the method of time series alignment.

7. The data for filling in Table 1 is calculated automatically, and this table is formed as the final result.

To compare the work of the Yolov7 Pose, Mediapipe Pose and Open Pose skeleton construction methods, a Python script was used that splits the video into frames using the OpenCV library. Then each frame was used as an input image of a neural network, the result of the network is a skeleton. Then the resulting skeleton was drawn over the frame using the OpenCV, scikit-learn and plotly libraries, and the results of the work of 3 networks on the same frame were combined into one image using the numpy library. After that, a visual analysis of the quality of the methods was carried out. The results are shown in Figure 4.

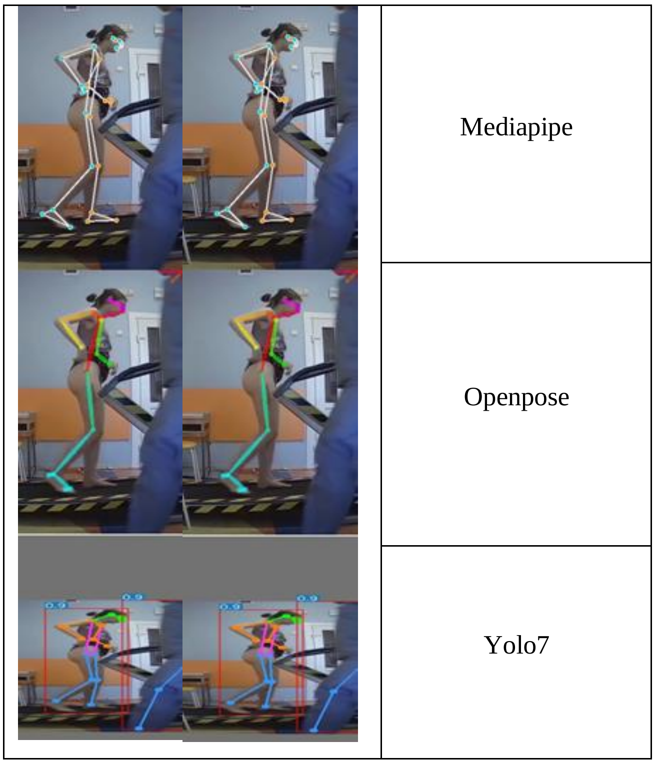

Fig. 4. Comparison of human skeleton construction methods

Figure 5 shows that Open Post may not recognize one of the legs, which is critical for gait analysis. In addition, part of the doctor's leg was detected as part of the patient's leg, which is also quite critical for analysis.

Mediapipe Pose and Yolov7 Pose produce equally stable results. Mediapipe Pose was chosen as the algorithm for constructing the skeleton in this task, because it recognizes more key points of the skeleton, specifically more points of the foot. This allows to perform better segmentation of the step, to determine the characteristics of the change in the position of the foot in different periods of the step.

V. Automation Of The Method Of Diagnostics Of Gait Disorders

For segmentation to determine the step period on each frame, it is necessary to be able to analyze the change functions of all available key points simultaneously. To do this, it is possible to use an algorithm of dynamic transformation of the timeline, which determines the optimal correspondence between sequences.

In Figure 5, there is some displacement between the extreme points of the graphs of the x-coordinates of the toe of the right foot and the angle of inclination of the foot to the treadmill. It is necessary to combine the extremes of the functions in order to remove local extremes that are not essential for the performing of the algorithm. To obtain better segmentation results, it is necessary to use data on changes in the coordinates of a larger number of body parts.

The dynamic transformation of the timeline (dtw) algorithm allows to find the optimal correspondence between sequences. In this paper we used its implementation metrics.dtw_variants-dtw_path module in the tslearn library.

n the standard implementation, this algorithm works for two time sequences.

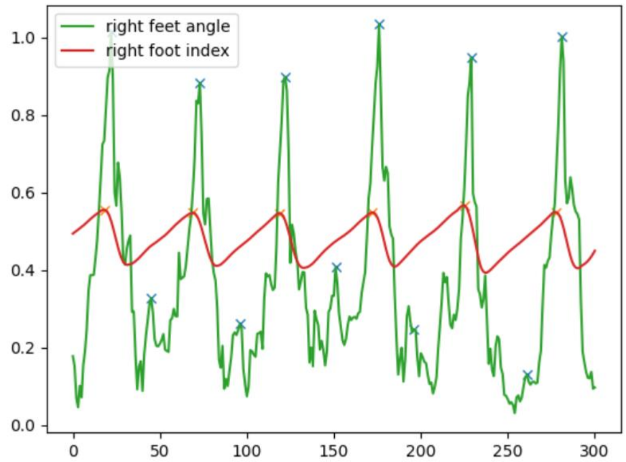

Fig. 5. Graph of the function of changing the angle of inclination of the foot and the x-coordinates of the toe of the right foot

To analyze more than two sequences, a modification of this algorithm is used (the time series alignment method), its implementation in the same module is dtw_barycenter_averaging_subgradient. The result of the algorithm for the functions of changing the x-coordinate of the left heel (blue line), changing the x-coordinate of the toe of the left foot (orange line) and changing the x-coordinate of the left knee (green line) is shown in Figure 6 (red line).

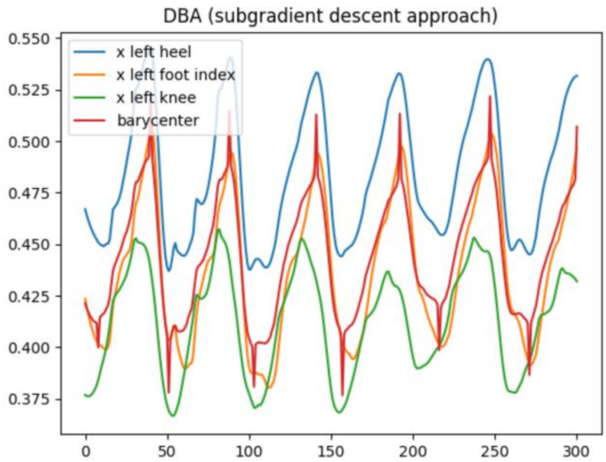

Fig. 6. The result of the work of the time series alignment method

This algorithm combines the extremes of functions, which allows to determine in which frame one step cycle was replaced by another.

There are two periods of double support in the step cycle. The end of the double support period is characterized by the angle between the feet of both legs and the treadmill. At the end of the double support period, the angle with the treadmill is minimal and close to zero for one foot, it is maximal for the second foot, since rolling occurs due to the transfer of body weight to the toe of the second foot (Figure 7). Also, the period of double support is characterized by minima of the x-coordinates of the heel and toe of the hind leg. After the double support period, the transfer period begins. This means that a person moves his hind leg forward, that is, increases the x-coordinate of all parts of the foot of this leg.

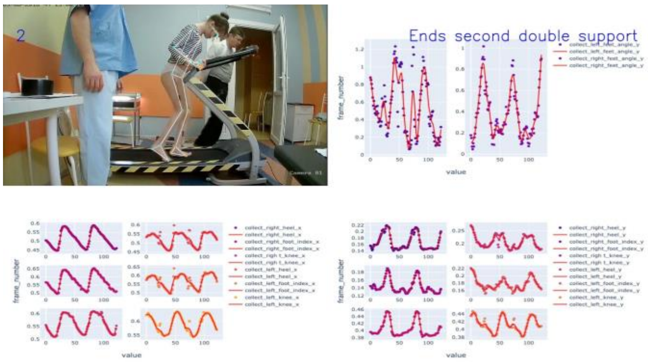

Fig. 7. The end of the double support period

The end of the period of the second double support differs only in that the back leg is the right, otherwise the logic is exactly the same as in the first period.

The end of the period of single support coincides with the beginning of the step of the left foot Using the same rule as for determining a new step from the right foot, only considering the coordinates of the left foot as an analysis, it is possible to determine the frame corresponding to the end of the period of single support. To get a more accurate result, it is possible to add an additional point (for example, in addition to analyzing the movement of the left heel, add an analysis of the movement of the toe of the left foot).

Thus, it is possible to find frames corresponding to the following events:

– the beginning of the step / the beginning of the period of the first double support / the end of the transfer period;

– the end of the first double support period/beginning of the single support period;

– the end of the single support period/start of the second double support period;

– the end of the second double support period/start of the transfer period;

Therefore, it is possible to completely segment the step and get all the necessary measurements.

The algorithm for automating the analysis of walking pathologies, implemented in Python, includes the following steps:

1. The resulting video fragment is divided into frames.

2. On each frame, a human skeleton is built using Mediapipe Pose.

3. The values of x, y coordinates are stored, discrete functions of coordinate changes from time are constructed.

4. The x, y coordinates of the following body parts are stored on each frame: left and right heel, left and right toe, left and right knee, left and right hip, angle between heel and treadmill in the sagittal plane, both legs.

5. The obtained discrete functions of limb coordinate changes are filtered using the Savitsky-Goley filter.

6. Step segmentation is performed:

i. the frame number from which the new step starts is equal to the frame number corresponding to the local maximum x-coordinates of the right heel;

ii. the end of the period of the first double support is equal to the frame number corresponding to the local minimum x-coordinates of the left heel, x-coordinates of the toe of the left foot, x-coordinates of the left knee, the angle between the left heel and the treadmill;

iii. the end of the single support period coincides with the beginning of the step of the left foot and is equal to the number of the frame corresponding to the local maximum x-coordinates of the left heel and x-coordinates of the left toe;

iv. the end of the period of the second double support is equal to the frame number corresponding to the local minimum x-coordinates of the right heel, x-coordinates of the toe of the right foot, x-coordinates of the right knee, the angle between the right heel and the treadmill;

7. The data necessary for diagnosis is automatically collected and a table is formed as the final result.

VI. Testing The Algorithm

To test the operation of the pipeline, a visualization in Python was developed using the plotly library. An example of the developed visualization is shown in Figure 8.

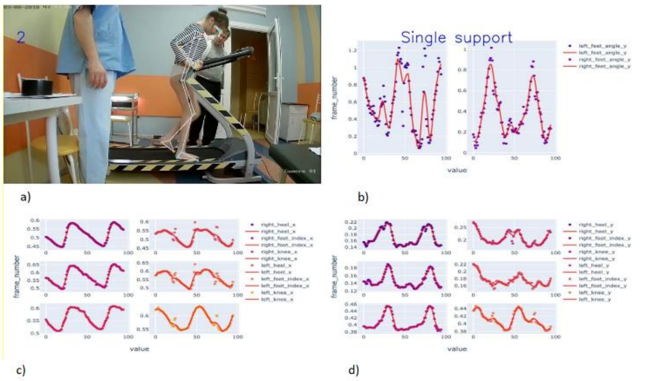

Fig. 8. Visualization of the algorithm: the image obtained from the camera (a); graphs of the functions of changing the angle of the foot to the treadmill (b); graphs of the functions of changing the x coordinates of different parts of the body from time (c); graphs of the functions of changing the y coordinates of different parts of the body from time (d)

This visualization helps to qualitatively evaluate the work of the algorithm, and determine the error of the segmentation algorithm. The number 2 in Figure 8a shows which right step a person is currently taking. The label "Single support" in Figure 8b defines the step period corresponding to this frame.

Figures 9-12 show graphs of functions for changing the position of key points. They allow to determine which period of the step cycle corresponds to a given frame, as well as changes in functions depending on the period of the step cycle. Based on this information, decisions about step segmentation are made.

The pipeline implemented in this paper automatically calculates the following step characteristics:

– duration of the step cycle;
– duration of the support phase;
– duration of the transfer phase;
– hip angle initial contact;

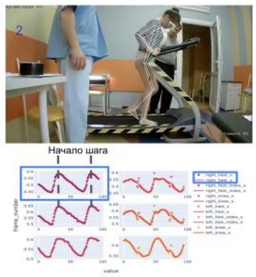

Fig. 9. The beginning of the step / the beginning of the period of the first double support / the end of the transfer period

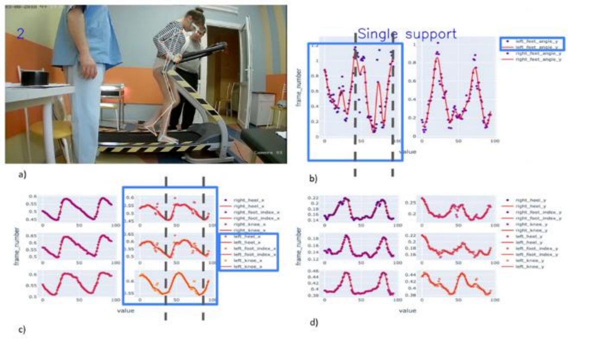

Fig. 10. End of the period of the first double support /beginning of the period of the second double support

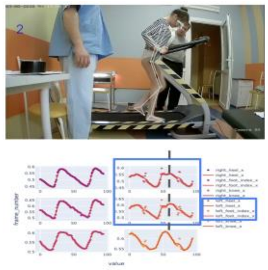

Fig. 11. End of the single support period/beginning of the second double support period

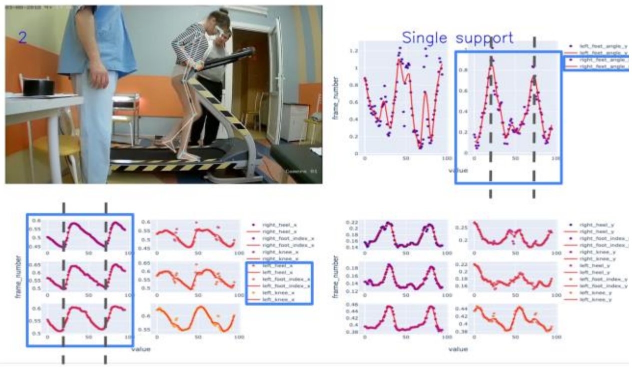

Fig. 12. End of the second double support period/beginning of the transfer period

– angle of the knee joint initial contact;
– angle of the ankle joint initial contact;
– hip angle double support period completion;
– knee joint angle double support period completion;
– ankle angle double support period completion;
– hip joint angle double support period start;

– knee joint angle double support period start;

– angle of the ankle joints the period of double support the beginning;

– hip joint angle transfer period vertical position of the tibia;

– angle of the knee joint transfer period vertical position of the tibia;

– angle of the ankle joint transfer period vertical position of the tibia.

Two video clips provided by the doctor were used as a dataset for testing. In total, 30 cycles of right steps were presented in both videos. Testing was divided into two parts: the accuracy of step segmentation and the accuracy of automatic calculation of the required characteristics. Manual data markup was done for both parts, with which the results obtained by the algorithm were then compared and the error was calculated. Table 2 shows the results of step segmentation testing. The values of the maximum average measurement error in frames are indicated. The maximum error is calculated as the difference between the frame that the person specified during the markup and the frame calculated by the system. The average error is the average of the absolute values of these differences.

Taking into account the fact that the fps of video fragments is known, Table 3 shows the results of testing the segmentation of the step with an error in seconds and Table 4 shows the results of angle testing.

TABLE II. STEP SEGMENTATION TEST RESULTS, ERROR VALUE IS SPECIFIED IN FRAMES

The beginning of the period of the first double support / the end of the transfer period, frame	The end of the period of the first double support / the beginning of the period of the single support, frame	End of the single support period/beginning of the second double support period, frame	The end of the period of the second double support / the beginning of the transfer period, frame
[-1, +2], 0.47	[-1, 0], 0.27	[-3, 0], 1.77	[-1, +1], 0.59

TABLE III. STEP SEGMENTATION TEST RESULTS, ERROR VALUES ARE SPECIFIED IN SECONDS

The beginning of the period of the first double support / the end of the transfer period, s	The end of the period of the first double support / the beginning of the period of the single support, s	End of the single support period/beginning of the second double support period, s	The end of the period of the second double support / the beginning of the transfer period, s
[-0.04, +0.08], 0.0188	[-0.04, 0], 0.108	[-0.12, 0], 0.0708	[-0.04, +0.04], 0.0236

TABLE IV. THE RESULTS OF TESTING AUTOMATIC CALCULATIONS, THE ERROR VALUES ARE INDICATED IN DEGREES

Key joint	Initial contact, degrees	Double support period (completion), degrees	Double support period (beginning), degrees	Transfer period (vertical position of the tibia), degrees
Hip	[-0.13º, +0.11º]	[-0.12º, +0.09º]	[-0.13º, +0.08º]	[-0.09º, +0.08º]
Knee	[-0.43º, +0.12º]	[-0.32º, +0.27º]	[-0.24º, +0.21º]	[-0.12º, +0.18º]
Ankle	[-0.89º, +0.3º]	[-0.67º, +0.23º]	[-0.87º, +0.05º]	[-0.76º, +0º]

As can be seen from the test results, the algorithm works quite stably, and the measurement errors are quite small, so doctors can use the obtained results. Comparison of the results of the duration of the step cycle, the support phase and the transfer phase are not given, since the measurement errors of these values are caused by the step segmentation error, which is shown in Table 1 and Table 2.

VII. CONCLUSION

This work is devoted to the study of the construction of the human skeleton and the automatic analysis of walking. The paper proposes a technology for diagnosing walking, which consists of the following main stages: building a human skeleton, determining the key points of the skeleton, registering the movement of key points, segmentation of walking and an algorithm for automated calculation of all step characteristics necessary for a doctor to make a diagnosis. Algorithms for step segmentation and extraction of all necessary characteristics have been developed. The technology is based on determining the frequency of step features, which allows you to quickly and efficiently determine all the necessary characteristics.

The implementation of this technology can be used by orthopedic doctors both as a separate program and for the development of other programs that will help speed up and improve the diagnosis of orthopedic diseases.

Further work will be derected at improving image segmentation for more accurate construction of the skeleton, as well as additional analysis of images in other projections to obtain additional characteristics.

This work is an important contribution to facilitating the analysis of the pathology of walking and significantly reduces the time required for doctors to diagnose diseases.

REFERENCES

[1] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei and Y. Sheikh, “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields” Available online: https://arxiv.org/abs/1812.08008. (accessed on 14.08.2023).

[2] H.-S. Fang, J. Li, H. Tang, C. Xu, H. Zhu, Y. Xiu, Y.-L. Li and C. Lu, “AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time”, Available online: https://arxiv.org/pdf/2211.03375.pdf. (accessed on 14.08.2023).

[3] R. A. Newcombe, Sh. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A.J. Davison, P. Kohi, J. Shotton and S. Hodges, “KinectFusion: Real-time dense surface mapping and tracking”, 10th IEEE International Symposium on Mixed and Augmented Reality. Basel, Switzerland, 2011, pp. 127–136.

[4] R. Atienza, “Pyramid U-Network for Skeleton Extraction From Shape Points”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2019, Long Beach, CA, USA, June 16-20, 2019, pp. 1–10.

[5] A. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang and M. Grundmann, “BlazePose: On-device Real-time Body Pose tracking”, Available online: https://arxiv.org/abs/2006.10204. (accessed on 14.08.2023).

[6] C.-Y. Wang, A. Bochkovskiy and H.-Y.M. Liao, “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors”, Available online: https://arxiv.org/abs/2207.02696. (accessed on 14.08.2023).

[7] D. Maji, S. Nagori, M. Mathew and D. Poddar, “YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation Using Object

Keypoint Similarity Loss", Available online: https://arxiv.org/abs/2204.06806. (accessed on 14.08.2023).

[8] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", Available online: https://arxiv.org/abs/1506.02640. (accessed on 14.08.2023).

[9] A. Bochkovskiy, C.-Y. Wang and H.-Y.M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection", Available online: https://arxiv.org/abs/2004.10934. (accessed on 14.08.2023).

[10] X. Zhu, S. Lyu, X. Wang and Q. Zhao "TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios", Available online: https://arxiv.org/abs/2108.11539. (accessed on 14.08.2023).

[11] V. Bazarevsky, Y. Kartynnik, A. Vakunov, K. Raveendran, M. Grundmann, "BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs", Available online: https://arxiv.org/abs/1907.05047. (accessed on 14.08.2023).

[12] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", Available online: https://arxiv.org/abs/1801.04381. (accessed on 14.08.2023).

[13] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications", Available online: https://arxiv.org/abs/1704.04861. (accessed on 14.08.2023).

[14] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, Ch-Y Fu and A.C. Berg, "SSD: Single Shot MultiBox Detector", Springer International Publishing Computer Vision (ECCV), vol. 9905, 2016, pp. 21–37.

[15] R. W. Schafer, "What Is a Savitzky-Golay Filter?", IEEE Signal Processing Magazine, vol. 28, no. 4, 2011, pp. 111-117.

[16] R. Girshick, J. Donahue, T. Darrell and J. Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation", Available online: https://arxiv.org/abs/1311.2524. (accessed on 14.08.2023).

[17] C.-Y. Wang, H.-Y.M. Liao and I.-H. Yeh, "Designing Network Design Strategies Through Gradient Path Analysis", Available online: https://arxiv.org/abs/2211.04800. (accessed on 14.08.2023).