Electronic Engineering and Informatics
G. Izat Rashed (Ed.)

doi:10.3233/ATDE240106

Visual Quantitative Evaluation of Berg Balance Scale for Stroke Patients

Zhihao WANG[a], Anming GU[b], Xiaofeng LU[a1]
[a]*School of Communication & Information Engineering, Shanghai University, No.99 Shangda Road, Shanghai, China*
[b]*Montverde Academy Shanghai, No.91 Jianhao Road, Shanghai, China*

Abstract. Stroke patients with hemiplegia are often accompanied by balance disability, which has a serious impact on their life, work and rehabilitation treatment. In order to estimate their recovery and formulate a rehabilitation plan, the Berg Balance Scale (BBS) is usually used to evaluate their balance disability. At present, BBS is often evaluated manually. But the human resources are tight in the hospital, and it is difficult for patients to participate regularly. In order to reduce the burden of both doctors and patients, this paper uses artificial intelligence algorithm to automatically score the test process of Berg Balance Scale. Firstly, the method uses HRNet (High-Resolution net) algorithm to obtain the 2D human bone information of patients in the video sequence. Then, MHFormer (Multi-Hypothesis transformer) network was used to predict the 3D bone information of patients. Finally, based on the features of human joint points, human motion recognition and quantitative scoring were carried out. The experimental results showed that the same effect could be achieved by using intelligent algorithm to score Berg scale.

Keywords. stroke, Berg balance scale, balance disability, human action recognition

1. Introduction

Stroke (stroke) refers to the cerebral ability defect syndrome (or acute cerebrovascular disease) caused by acute cerebral circulation disability. It has the characteristics of high incidence, high mortality, high disability rate and high recurrence rate. Hemiplegia is the most common and serious symptom in stroke patients, which is one of the main factors affecting their activities of daily living [1].

Balance ability is one of the important abilities of human body and an important indicator of human health. Balance disability ranks first among stroke related diseases (about 87.5%) [2]. It seriously affects patients' quality of life, and greatly increases the risk of falls, posing a serious threat to their human safety and health [3]. Balance ability assessment is an important part of balance rehabilitation. Effective balance assessment is necessary for evaluating rehabilitation efficacy and formulating an efficient balanced rehabilitation plan.

[1] Corresponding author: Xiaofeng LU, School of Communication & Information Engineering, Shanghai University, e-mail: luxiaofeng@shu.edu.cn

Table 1. Used by doctors to record Berg scale scores for patients

Scoring table of berg balance scale			
Item name	Score(0-4)	Item name	Score(0-4)
1.sitting to standing		8.reaching forward with arms	
2.standing unsupported		9.retrieving object from floor	
3.sitting unsupported		10.turning to look behind,	
4.standing to sitting		11.turning 360 degrees	
5.transfers		12.placing foot on stool	
6.standing with eyes closed		13.standing with one foot in front	
7.standing with feet together		14.standing on one foot	
Total score(0-56)			

At present, the scale method is the gold standard for balance ability assessment. The Berg balance scale(BBS) is the most commonly used balance ability assessment scale. Its reliability and validity have been proved by many studies [4][5]. BBS is a commonly used scale to evaluate the balance ability of the elderly and patients with movement disability.

As shown in table 1, the assessment includes 14 items: sitting to standing, standing unsupported, sitting unsupported, standing to sitting, transfers, standing with eyes closed, standing with feet together, reaching forward with outstretched arms, retrieving object from floor, turning to look behind, turning 360 degrees, placing alternate foot on stool, standing with one foot in front, standing on one foot. Each item is scored 0-4 points, with a total score of 56 points (approximately 20 minutes to complete the test). All items should be evaluated by the same therapist.

We use methods of computer vision to obtain real-time patient movements based on the camera, analyze and complete the automatic evaluation of the Berg scale. In this way, patients can complete the assessment of balance ability at home, which is greatly convenient for both doctors and patients.

2. Algorithm Design

Recognizing and scoring the patient's test action in the video is a task of human action recognition. It usually requires two steps: the first step is to use the human pose estimation algorithm to recognize the bone joint points, and the second step is to extract features from the bone information, then use the classifier to classify the action.

We first use HRNet algorithm to extract the 2D human pose data of the characters in the monocular video, and then uses MHFormer algorithm to extract the 3D human pose data. Then we use the traditional method based on the characteristics of human joints to extract the characteristics of human action recognition, and design a classifier according to the rating requirements of Berg scale.

2.1 2D human pose extraction based on HRNet algorithm

HRNet is a new type of deep learning network, which can effectively solve the problem of image semantic segmentation. It is a human pose estimation model proposed by the University of science and technology of China and Microsoft Research Asia in 2019 for the 2D human pose estimation (human pose estimation or key-points detection) task. It performs well in the three tasks of key-point detection, pose estimation and multi person pose estimation in the coco dataset [6].

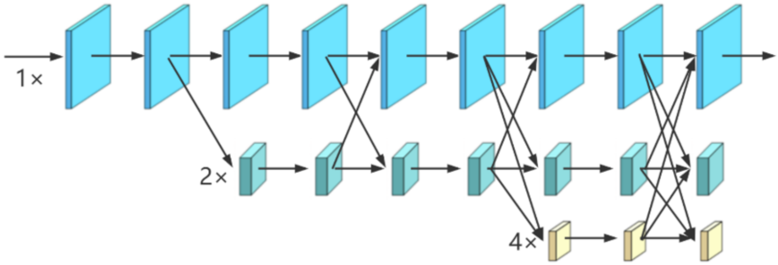

Figure 1. HRNet Network Architecture

HRNet uses high-resolution subnets in the first stage, gradually increasing the number of high-resolution to low-resolution subnets to form more stages, and connecting these subnets with different resolutions in a parallel manner. Different resolution representations can repeatedly obtain information from representations of other resolutions, thus generating rich high-resolution representations, as shown in Figure 1. This way, the predicted key-point heat map is more accurate.

Figure 2. Schematic diagram of key-points of 2D human body

We input the image to be detected into the feature extraction network HRNetV2-W32. After the first three stages, we get four different scale feature maps in stage 4.The four feature maps are upsampled by bilinear interpolation and convolution processed by $1\times1\times128$、$1\times1\times128$、$1\times1\times256$、$1\times1\times256$ convolution kernel .Then we splice the four convolution processed feature maps to get the feature map with the size of the input image.

As shown in Figure 2, it is a schematic diagram of key-points of two-dimensional human body. It is trained on the dataset COCO and Human3.6M with coco_h36m model HRNet. It consists of 17 nodes, namely 0-hip, 1-hip right, 2-right knee, 3-right ankle, 4-hip left, 5-left knee, 6-left ankle, 7-chest, 8-neck, 9-jaw, 10-head, 11-left shoulder, 12-left elbow, 13-left wrist, 14-right shoulder, 15-right elbow, and 16-right wrist.

2.2 3D human pose extraction based on MHFormer algorithm

Due to the ambiguity and self-occlusion of depth maps, estimating 3D human pose based on monocular videos is a challenging task. To address these issues, most methods focus on exploring the relationship between spatial and temporal domains. Due to the lack of depth information, there are multiple feasible solutions for 2D to 3D enhancement of monocular videos. In recent years, there has been a lot of research to address this issue. They typically rely on one-to-many mapping, adding multiple output headers to existing architectures through a shared feature extractor, without establishing relationships between different assumed features.

The MHFormer algorithm uses a more reasonable method: it firstly performs a one-to-many mapping, then performs a many-to-one mapping and various intermediate assumptions, thereby enriching the diversity features and ultimately synthesizing a better 3D human pose estimation.

MHFormer has built a three stage framework:1) Generating multiple initial hypothesis representations; 2) Establishing a self-hypothesis communication model, merge multiple hypotheses into a convergent representation, and then divide it into several divergent hypotheses; 3) Learning cross hypothesis communication, aggregate multiple hypothesis features, and synthesize the final 3D pose. Through the above processing, the final representation was enhanced and the accuracy of the synthesized pose was improved [7].

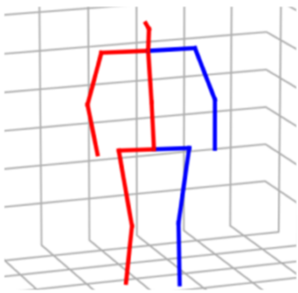

Figure 3. MHFormer Algorithm for 3D Human key-Point Extraction

In Figure 3, the MHFormer algorithm is used to predict the 3D human key-points based on the previously extracted 2D bone data. These nodes are corresponding to the 2D key-points one by one.

2.3 Feature Extraction and Quantitative Scoring

The current mainstream method for designing classifier is to use deep learning to train models from datasets. However, due to the lack of a suitable dataset for many actions of the Berg scale, such as bending down to pick up objects or alternating feet on stools. We adopt the traditional method of manually designing classifiers based on prior knowledge. Firstly, we design a classifier based on the scoring requirements of the Berg scale, and then design the features to be extracted according to the classification criteria of the classifier.

The design of the classifier mainly starts from two perspectives: one is to inspect whether the action is completed. If it is not completed, it will be directly assigned a score of 0. The other is to refine the score based on the quality of the action completion.

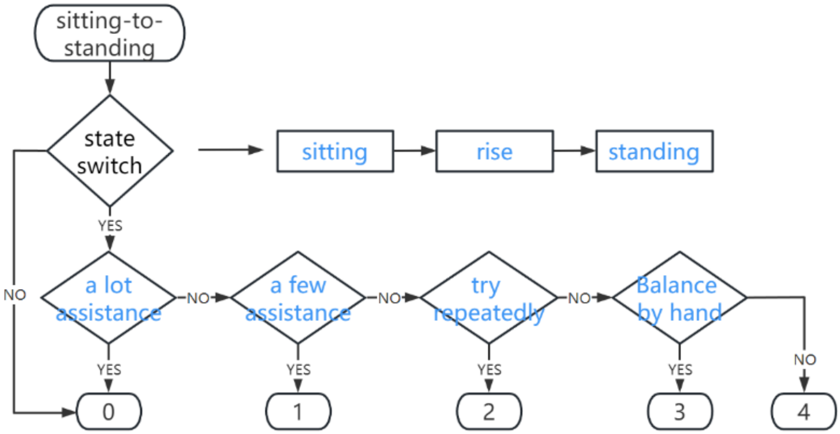

Figure 4. Schematic diagram of the design of the station arrival classifier for item 1

As shown in Figure 4, we take item 1 as an example. Its action requirement is to stand up, and try not to support with hands. Score and scoring type during manual evaluation are as follows:

(4) Able to stand and maintain balance without the need for hands

(3) Able to stand and maintain balance with hands

(2) You can stand and maintain balance with your hands, but you need to try several times

(1) Standing or maintaining balance requires a small amount of assistance from others

(0) Standing or maintaining balance requires a lot of assistance from others

The sequential switching between three states(sitting, rising, and standing) indicates completion of the action. We use the line formed by the two feature points of the left hip and knee in the 3D bone data to represent the left thigh. Then we calculate the angle between the line and the vertical line. The complementary angle of the angle formed by two lines is the angle between the left thigh and the horizontal plane. The angle determines whether the patient is sitting, rising, or standing. The angle less than 15 degrees refers to sitting position. The angle greater than 45 degrees refers to the standing position. The angle in between refers to the rising position. If the patient has not completed the transition from sitting position to rising position and then to standing position, it indicates that the action has not been completed.This test will be assigned a score of 0.

If the patient completes the action, we need to further confirm the following points:

1.the independence of action

When using the HRNet algorithm to extract 2D bone data, a corresponding set of data is generated for each character in the screen. The number of groups in the bone data can be used to determine the number of people in the screen. Multiple people indicate the presence of assistance.

2. the level of assistance

If the abscissas of both wrists feature points of the helper are both less than or greater than the abscissa of the tester's crotch feature points, it indicates that the helper is a unilateral arm assist, with a relatively light degree of assistance. If not, it indicates that the helper is an encircling assist, with a relatively heavy degree of assistance.

3.the fluency of action

From frame 6, we start to calculate the angle between the left thigh and the horizontal plane in this frame and the average angle between the left thigh and the horizontal plane in the previous 5 frames of this frame. If the result of this frame is less than the average value of the previous 5 frames, it indicates that there is reciprocating movement when standing up and the action is not smooth).

4.the difficulty of maintaining balance

We first calculate the distance between the feature points of the two wrists and the distance between the feature points of the two shoulders, and use the two distances as a quotient. If the result is greater than the threshold, it indicates that the hands are stretched wider to maintain balance. This indicates that maintaining balance is a bit difficult for patients.

3. Analysis of Experimental Results

The experimental equipment is the Lenovo notebook computer yoga14s2020 with AMD ryzen7 4800u processor. The video resolution is 640 * 480, and there is no need for a graphics card device during operation.

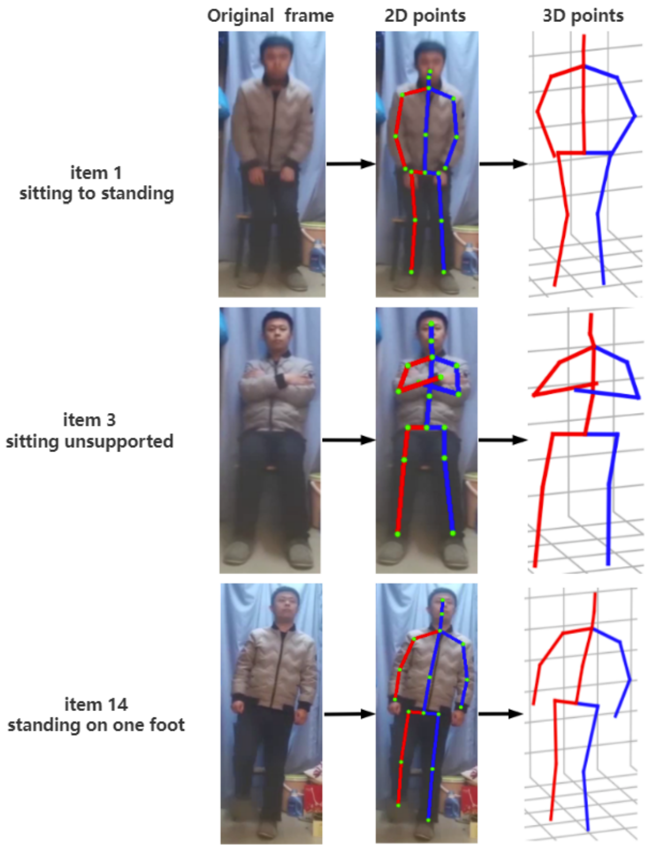

Figure 5. Schematic diagram of original video frame and key-points of human body

As shown in Figure 5, the scoring algorithm can accurately identify the key-points of human bones in two and three dimensions for patient images in different poses.

In this experiment, we used both manual and algorithmic methods to score the three groups of videos separately. The first group is the video in of people without balance disability, and the second and third groups are the videos simulating the balance disability under different items. The algorithm will first extract the 2D and 3D bone data of each frame image in the video, and then input the bone data into the classifier of the corresponding item according to the item number marked on the video. The classifier will classify the action quality like item 1 in the example above, and output the score. When scored manually, the three groups of videos are strictly scored by the same therapist.

Table 2. The results of scoring three groups of videos with manual and algorithmic methods separately

Comparison table of experimental results																
Patient No.	Item No.	01	02	03	04	05	06	07	08	09	10	11	12	13	14	Total score
1	Manual	4	4	4	4	4	4	4	4	4	4	4	4	4	4	56
	Algorithm	4	4	4	4	4	4	4	4	4	4	4	4	4	4	56
2	Manual	3	2	2	3	1	2	2	3	2	1	2	3	3	3	32
	Algorithm	3	2	2	3	1	2	2	2	2	1	2	3	2	3	30
3	Manual	2	3	3	1	3	3	1	2	3	3	3	2	2	2	33
	Algorithm	2	3	3	1	3	3	1	2	2	3	3	2	2	2	32

As shown in table 2, the results of the evaluation algorithm in the first group are completely consistent with the scores given by the manual method. However, in the second group, the evaluation algorithm results for item 8 (reaching forward with outstretched arms), item 13 (standing with one foot in front), and item 9 (retrieving object from floor) in the third group did not match the manual scores. This is due to the error in the predicted depth information when using the MHFormer algorithm to extract the key-points of the 3D human body. It leads to the deviation of the results when judging the patient's bending angle, arm extension distance, and the distance between the front and rear feet, and the classification accuracy is slightly affected.

4. Conclusion

The traditional way of manually evaluating Berg scale requires the participation of doctors to assist and score. It makes many patients who have rehabilitation training at home unable to regularly carry out balance assessment, thus affecting the recovery progress of balance ability. In this paper, we use computer vision to extract the 3D bone information of patients in the test video, and then recognize the completion of the test action based on the characteristics of joints, which effectively realizes the automatic evaluation of Berg scale. Finally, based on the features of human joint points, human motion recognition and quantitative scoring were carried out. The experimental results showed that the same effect could be achieved by using intelligent algorithm to score Berg scale. Stroke patients with balance disability can use this algorithm to evaluate their balance ability and carry out reasonable recovery treatment.

References

[1] Chen Xiu Xiu,Wu Qing Wen,Guo Zi Meng, Liu Guang Tian, Cui Ying. Effect of mirror therapy on lower limb motor ability, activities of daily living and balance ability of stroke patients [J]. Chinese Journal of rehabilitation medicine, 2019,34(05):539-543+55012

[2] Lendraitienė E, Tamošauskaitė A, Petruševičienė D, et al. Balance evaluation techniques and physical therapy in post-stroke patients: a literature review [J]. Neurol Neurochir Pol, 2017, 51(1): 92-100.

[3] Van Duijnhoven HJ, Heeren A, Peters MA, et al. Effects of exercise therapy on balance capacity in chronic stroke: systematic review and meta-analysis [J]. Stroke,2016, 47(10): 2603 - 2610.

[4] Jin Dongmei, Yan tiebin, Zeng Haihui. Validity and reliability of Berg Balance Scale [J]. Chinese Journal of rehabilitation medicine,2003(01):24-26.

[5] Alghadir AH, Al-Eisa ES, Anwer S, et al. Reliability, validity, and responsiveness of three scales for measuring balance in patients with chronic stroke [J]. BMC Neurol, 2018, 18(1): 141.

[6] Sun Ke,Xiao Bin,Llu Dong,et al.Deep highresolution representation learning for human poseestimation [C]//Proceedings of 2019 IEEE/CVFConference on Computer Vision and Pattern Recognition (CVPR) .Long Beach,CA: IEEE,2019: 5686-5696.

[7] Li W, Liu H, Tang H, et al.MHFormer:Multi-Hypothesis Transformer for 3D Human Pose Estimation [J]. 2021.