

Introduction

Accurate identification of cerebral palsy (CP) plays an important role in the initial diagnosis and future treatment evaluation. A very common method for identifying CP is through gait analysis. However, current state-of-the-art gait analysis systems require not only extensive equipment but also extensive use of markers and sensor attachments on patients. As a result, state-of-the-art systems pose two issues. From a patient standpoint, accessibility to these systems can potentially be cumbersome for those who live farther away from these systems. In smaller hospitals and practices, these state-of-the-art systems could also be financially out of reach.

Our proposed method takes individual videos as input, estimates their 3D body pose using our proposed DNN-based model, and then analyzed by our classifier, determining whether the patient is healthy or diagnosed with CP. We trained our model from a dataset consisting of 6 healthy children and 6 CP children, with our experimental results showing an overall classification accuracy of 91.7%. This study serves as a starting point for the development of more robust tools for automatic classification of gait impairments and as a basis for future Deep Learning applications in clinical gait analysis.

Methods

Data Acquisition

Our dataset includes walking pattern records of 6 healthy children and 6 CP children. Subjects were asked to walk on a treadmill for about one minute with a digital camera recording their gait pattern and a synchronized motion capture system directly measuring their body movement. Digital camera was located on the sagittal plane of the subjects and had 480x640 pixels resolution. 8 Reflective markers were attached to the neck, chest, left/right hips, left/right knees, and left/right ankles, which were traced by a motion capture system with a sampling rate of 100 Hz.

Data Pre-processing:

To prepare the data, we first extract images from each video frame. Images are adjusted to 256x256 pixels and are cropped such that the subject is located at the center. 3D joints annotations, provided by a marker-based motion capture system, are normalized from zero to one. 2D joints annotations are calculated by registering 3D joints annotation into image coordinates system. After pre-processing, the data structure consists of the cropped images, corresponding 2D joints annotation, and normalized 3D joints annotation.

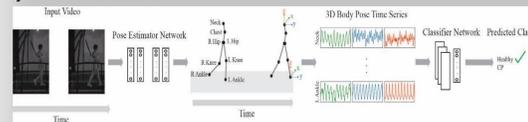


Figure 1: Overview of the proposed system

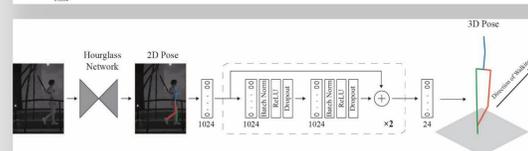


Figure 2: Architecture of the "Pose Estimator" network

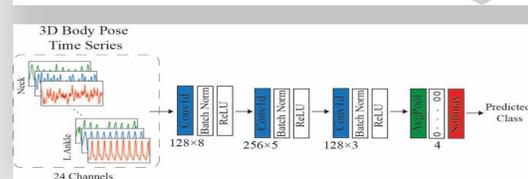


Figure 3: Network of the "Classifier" network

Objectives

- 1) Increase the accessibility to CP analysis tools for both patients and physicians.
- 2) Reduce the cost of equipment while providing the possibility for remote physician appointments using Deep Neural Networks (DNNs)

Results

3D Pose Estimation Results:

The ground truth (obtained from the marker-based motion capture system) and estimated 3D joint coordinates are quantitatively compared by using 3D pose error. 3D pose error is calculated based on the average of Euclidean distance between estimated 3D joints coordinates and corresponding ground-truth data for all joints. Table 1 shows the 3D pose error for each subject and group separately. Averaged 3D pose errors is 39.19±8.71 mm on the whole dataset and no significant difference is observed between healthy and CP group subjects. For qualitative results, we have provided representative 3D poses predicted by our proposed method in Figure 4.

Figure 4: Qualitative results of 3D pose estimation

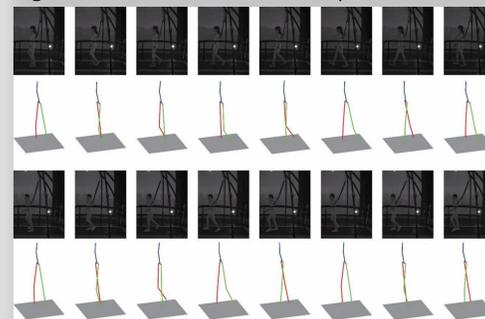


Table 1: Average 3D pose error (mm) for each subject group separately

Group	Subject	Mean (mm)	SD (mm)
CP	1	30.43	4.92
	2	51.15	7.09
	3	39.44	5.13
	4	44.92	8.65
	5	33.36	5.88
	6	34.45	5.05
Healthy	1	31.17	7.71
	2	35.95	8.36
	3	45.37	9.76
	4	35.27	6.44
	5	45.92	7.51
	6	42.86	19.04

Gait Classification Results:

In order to assess the performance of classifier, accuracy, sensitivity, and precision measures are used. Recall represents the percentage of actual positives cases that are identified correctly, while precision represents the percentage of positive identifications that are actually correct. Accuracy on the other hand, shows the overall percentage of cases that are classified correctly. On our data set, we observed 91.7% accuracy, and sensitivity and precision were %83.3 and %100.0, respectively. Table 2 shows the confusion matrix for each class. As shown in table 2, no false positive and only one false negative case happened, where one of the CP children was misclassified as a healthy child.

Table 2: Confusion matrix for gait classification from estimated 3D pose time series

		Classification Output	
		Healthy	CP
Actual Class	Healthy	6	0
	CP	1	5

Discussion

Findings: The proposed system removes the requirement of complex equipment and large laboratory space and does not require domain medical knowledge for feature engineering. Results showed that the proposed system can detect a CP case with high confidence and safety (rare false negative) from only one digital camera and demonstrates the potential of the proposed system for in-home gait monitoring of patients.

Compared to the previous literature, the present study is novel in two aspects. Firstly, the proposed DNN could achieve similar or even higher classification accuracy without implementing a human-crafted feature selection and by only feeding raw 3D body pose to the network due to the ability of the DNNs to learn semantic and high-level features from time-series autonomously make predictions. Secondly, the current method is unique because it is applied on pathology gaits of patients compared to previous studies that utilized their methods on healthy individuals with normal walking pattern or on abnormal gait simulated by healthy people.

Analyzing results revealed that the healthy group achieved better classification accuracy. Except for one false negative, all the subjects were classified correctly. However, due to the limited size of the available dataset, it resulted in a 16.7% penalty to the classification accuracy of the CP group. No correlation was observed between the classification accuracy and pose estimation accuracy, but because of the small size of the dataset, a generalization of these results should be made with caution.

Future Vision: To our knowledge, this paper is the first to use a deep neural network to classify CP gait from the video. The goal of this research is to provide an ambient assisted living tool for a constant gait monitoring of patients in the domestic environment by taking advantage of DNNs. With the decrease in equipment and setup complexity of these systems, we hope to see possibilities of remote assessments to reduce travel to clinicians and the deployment of gait analysis systems both in homes and other clinics. With enough data, we hope to provide a basis for creating an objective rating system to calculate CP severity quantitatively.

References

- Barton, J. and A. Lees, 1997. "An application of neural networks for distinguishing gait patterns on the basis of hip-knee joint angle diagrams." *Gait & posture* 5(1): 28-33.
- Martinez, J., R. Hossain, J. Romero and J. J. Little, 2017. A simple yet effective baseline for 3d human pose estimation. *Proceedings of the IEEE International Conference on Computer Vision*.
- Newell, A., K. Yang and J. Deng, 2016. Stacked hourglass networks for human pose estimation. *European Conference on Computer Vision*, Springer.

Presented at the Aresty Research Center Undergraduate Research Symposium April 24th, 2020
Supported by NSF (IIS 1703883)
Please send correspondence to Kang Li (kl419@soe.rutgers.edu)