Automatic Evaluation System for Online Physical Education

Kazumoto Tanaka
Kindai University

Learning management system (LMS) is a software application that supports e-learning. However, there are some classes in which data management and automatic scoring by LMS are difficult. Dynamic physical education is one such class. Although a teacher can observe a student's physical performance via a remote video camera, the evaluation is done visually. Therefore, data analysis and data management cannot be automated. To address this problem, we have developed an automatic evaluation system for students' exercise demonstration videos. The proposed system detects the human body from the video and quantifies students' movement to automatically evaluate the performance of the exercise.

Keywords: automatic evaluation system, deep neural network, key point detection, physical performance, online physical education

INTRODUCTION

The coronavirus disease (COVID-19) has created a need for more extensive distance education. This has led to a distinctive rise in e-learning. Learning management system (LMS) is a software application that supports e-learning. Teachers can use an LMS to efficiently manage the student learning processes. However, there are some classes in which data management and automatic scoring by LMS are difficult. Dynamic physical education is one such class.

Most previous studies on e-learning for physical education have involved wearable sensors (e.g. wearable accelerometers) and applications to monitor students' body movement (Yang, et al., 2020). However, it is difficult to evaluate their movement correctly. For example, it is difficult to determine whether the movement measured by the sensor is due to correct body movement, or whether the sensor was just shaken while sitting down.

Another approach has utilized video camera image obtained from student's smartphone to check their motion form and performance (Steinberg, et al., 2020). Although a teacher can observe a student's physical performance via a remote video camera, and record it as digital data, the evaluation is done visually. Therefore, data analysis and data management cannot be automated. Furthermore, the teacher's visual acuity is heavily burdened when multiple videos which are sent from students are being visually evaluated on a computer display.

To address this problem, we have developed an automatic evaluation system for students' exercise demonstration videos (Tanaka, 2020). The proposed system detects the human body from the video and quantifies students' movement to automatically evaluate the performance of the exercise. The motion detector was developed based on a deep neural network (DNN) framework. Experiments were carried out to evaluate the system.

AUTOMATIC EVALUATION SYSTEM

The proposed system detects the position of body parts (e.g. head) to identify the student's pose from the video images taken by their smartphone during e-learning for dynamic physical education. This paper refers to the center position of a body part as a "key point." The student's motion can be recognized by using the key points in the time series. In this research, Mask R-CNN (He, et al., 2017) has been employed for the detector of the key points. Mask R-CNN is a framework for object detection and image segmentation based on a deep neural network.

We assume that a physical education teacher uses a web-based video conferencing tool such as Zoom app to observe students' exercise. First, our system captures the desktop screen where the Zoom is running (See Figure 1, in this figure, students are doing push-ups) and records the captured images during students' exercise. After the recording, the system determines the human regions and key points based on Mask R-CNN using the recorded images. The input image to Mask R-CNN is an image with unnecessary parts for the detection in the Zoom window cut off to reduce the processing time for the key point detection. Figure 2 shows a sample result of the key point detection. In this figure, the key points of both shoulders and both elbows were detected to evaluate student's push-ups.

After all the images are processed, displacement of key point position and angles between key points are calculated to evaluate number of movements, motion speed, and pose accuracy in the exercise. The system arranges the results according to the order of student video window and records them in a csv file. A sample of the results is shown in Figure 3. The figure shows the number of times for push-ups and squat.

FIGURE 1
DESKTOP SCREEN IMAGE OF THE ZOOM INTERFACE

FIGURE 2
RESULT OF KEY POINT DETECTION

FIGURE 3
A RESULT OF THE ANALYSIS

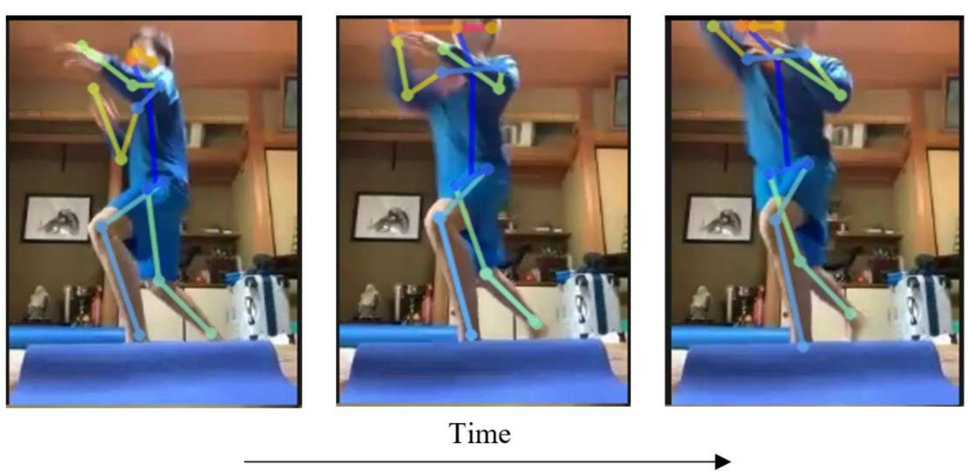
	A	B	C	D	E
1	student	push-up	squat	date 2020-8-20	
2	R1C1	45	90		
3	R1C3	41	98		
4	R2C1	40	87		
5	R2C2	56	91		
6	R2C3	42	90		
7	R3C1	45	92		
8	R3C2	42	89		
9					

EXPERIMENT RESULTS

As a preliminary experimental evaluation of the proposed system, we asked a university karate team coach for using the system to evaluate the usefulness of the system for e-learning of physical training. The coach used Zoom to remotely instruct seven members of the team, and the system was employed to evaluate the number of times for push-ups and squat. As a result, it was found that the system can count the number of times for simple exercise such as push-ups and generate the csv file successfully (See Figure 3).

For further evaluation, key point detection for karate techniques such as punches and kicks were tried. The members were instructed to set their body orientation appropriately so that the motion of the techniques could be observed by the coach. Figure 4 shows a sample of the detection results for punch motion by the right hand. Figure 5 shows the horizontal displacement of the right hand and the centre of both shoulders in this motion. Image frame 1 is the start of the motion, and the displacement at this frame was set to zero. It can be seen that the right thrust extends greatly to frame 4, and then the whole body moves forward. This shows the characteristics of karate thrust, and it is considered possible to automatically analyse the punch motion as in this example.

FIGURE 4
SAMPLE OF RESULTS OF KEY POINT DETECTION FOR PUNCH MOTION



On the other hand, key point detection for kick motion often failed as shown in Figure 6. The right image in the figure shows the error result. In this image, the kicking leg is greatly raised and the posture is standing on one leg. In many other kick images, detection failed in images with similar postures. The DNN model used in the detector may be lacking in learning such posture. In order to automatically evaluate various physical education exercises, it may be necessary to add learning results of special postures to the model. Clarifying this is a future task.

On average, the process required approximately 130 ms per image for key points detection on a PC (Windows OS, 3.3 GHz Intel Core i9 CPU, NVIDIA GeForce GTX 1080 Ti GPU). It can be said that the processing time is on a practical level.

FIGURE 5
DISPLACEMENT OF RIGHT HAND AND CENTER OF SHOULDERS IN PUNCH MOTION

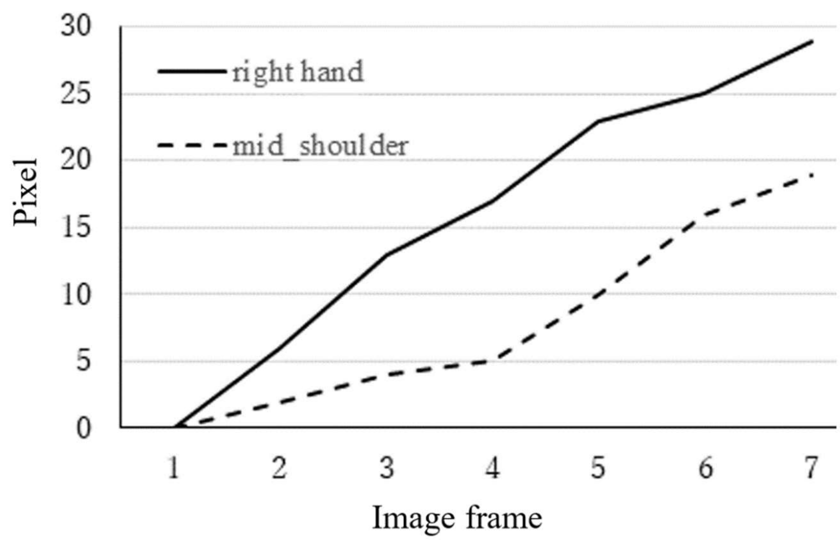
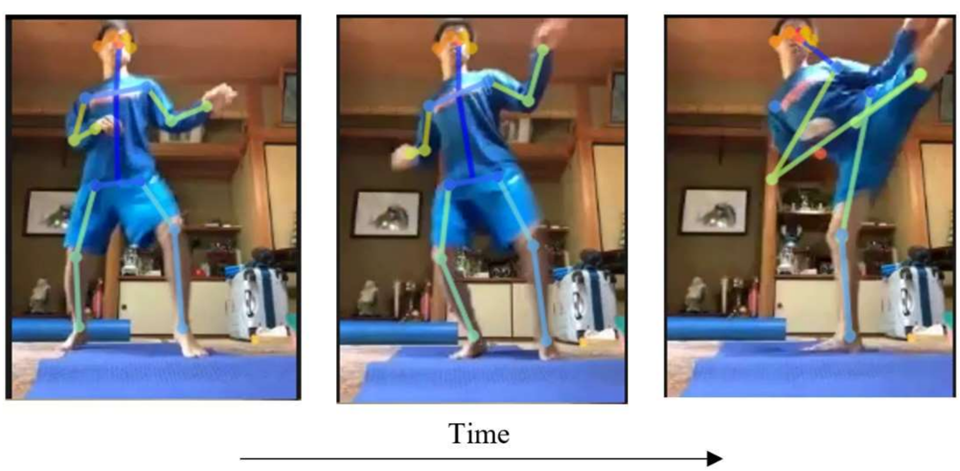


FIGURE 6
SAMPLE OF RESULTS OF KEY POINT DETECTION FOR KICK MOTION



CONCLUSIONS

We have developed an automatic evaluation system for students' exercise demonstration videos that can be used as an LMS on an e-learning for physical education. The system detects the position of human body parts (key points) from the web-based video sent from students by using the key point detector based on DNN (Mask R-CNN). The system can calculate displacement of key point position and angles between key points.

The system was evaluated using a video of a simple exercise such as push-ups and karate techniques. The result for the simple exercise showed that the system can count the number of times. However, a problem was revealed that the key point detection failed when the system was applied for kick motion in karate technique. Therefore, the application of the current system is limited to the evaluation of simple exercises.

As a future work, we will focus on improving the robustness of the key point detection and verifying the effectiveness of the system in detail.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Number 18K02876.

REFERENCES

He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017). Mask R-CNN. In *Proc. of IEEE International Conference on Computer Vision (ICCV2017)*, pp. 2980–2988.

Steinberg, C., Zühlke, M., Bindel, T., & Jenett, F. (2020). Aesthetic education revised: a contribution to mobile learning in physical education. *German Journal of Exercise and Sport Research*, *50*, 92-101.

Tanaka, K. (2020). An Automatic Evaluation System for Students' Exercise on E-kearning for Physical Education. In *Proc. of 13th Annual International Conference of Education, Research and Innovation (ICERI2020)*.

Yang, Q-F., Hwang, G-J., & Sung, H-Y. (2020). Trends and research issues of mobile learning studies in physical education: A review of academic journal publications. *Interactive Learning Environments*, *28*(4), 419-437.