data

IntelliRehabDS (IRDS)—A Dataset of Physical Rehabilitation Movements

Alina Miron [1,†]📷 , **Noureddin Sadawi** [1,2,†], **Waidah Ismail** [3,4,5]📷 , **Hafez Hussain** [6] **and Crina Grosan** [1,*]

1   Department of Computer Science, Brunel University London, Uxbridge UB8 3PH, UK; alina.miron@brunel.ac.uk (A.M.); noureddin.sadawi@gmail.com (N.S.)
2   Biotechnology Research Center, Tripoli TIP3644, Libya
3   International Halal and Fatwa Center, Universiti Sains Islam, Nilai 71800, Malaysia; waidah@usim.edu.my
4   Faculty of Science and Technology, Universiti Sains Islam, Nilai 71800, Malaysia
5   Information System Study Program, Universitas Airlangga, Indonesia Kampus C, Surabaya, Jawa Timur 60115, Indonesia
6   Perkeso Rehabilitation Centre, Bemban, Melaka 75450, Malaysia; hafez.hussain@rehabmalaysia.com
*   Correspondence: crina.grosan@brunel.ac.uk
†   These authors contributed equally to this work.

Abstract: In this article, we present a dataset that comprises different physical rehabilitation movements. The dataset was captured as part of a research project intended to provide automatic feedback on the execution of rehabilitation exercises, even in the absence of a physiotherapist. A Kinect motion sensor camera was used to record gestures. The dataset contains repetitions of nine gestures performed by 29 subjects, out of which 15 were patients and 14 were healthy controls. The data are presented in an easily accessible format, provided as 3D coordinates of 25 body joints along with the corresponding depth map for each frame. Each movement was annotated with the gesture type, the position of the person performing the gesture (sitting or standing) as well as a correctness label. The data are publicly available and were released with to provide a comprehensive dataset that can be used for assessing the performance of different patients while performing simple movements in a rehabilitation setting and for comparing these movements with a control group of healthy individuals.

1. Introduction

The assessment of human motion quality has applications in several domains: sports movement optimisation, range-of-motion estimation, and movement quality assessment in order to make a diagnostic assessment or for use as a tool in physical therapy and rehabilitation settings. Experts, such as coaches, physiotherapists and doctors, have been trained extensively to recognise what makes a certain motion *correct*. Building an automatic system for this task is not an easy endeavour, having to deal with a wide diversity of movements, human body capabilities and a certain degree of subjectivity. Kinect [1] (or other similar devices) camera-based sensor exercises are very common nowadays as they do not require any physical interaction with the subject [2–4].

With this in mind and with (still) the lack of available datasets containing gestures recorded from a correctness perspective, we created a platform and implemented it in a rehab centre. It was there that we collected real data from patients undergoing rehabilitation. Most of the existing datasets were built using healthy subjects who were asked to perform

both correct and incorrect (on purpose) executions [5–7]. The incorrect executions are often simulated by healthy people. In contrast, the data from our patients contain both correct and incorrect executions of gestures, both performed in a natural and free way.

We believe that the repository we made available is an excellent resource for the research community, especially for those working on software methods for motion quality assessment. In particular, the machine learning community will directly benefit from it as a platform for developing, improving and applying methods not only for gesture classification but also for gesture quality assessment (in terms of correctness) [8–10].

2. Related Work

There have been few initiatives about how to approach the problem of automatically assessing the level of correctness of a movement. Some of the ideas rely on using sensors attached to the body. In [11], the authors gathered a dataset using five sensor devices attached to the ankles, wrists and chest in order to record six exercises performed by 27 athletes and to label the data with a qualitative rating from one to five.

The Toronto Rehab Stroke Pose (TRSP) dataset [12] consists of 3D human pose estimates of stroke patients and healthy subjects who performed a set of movements using a stroke rehabilitation robot. The data recorded were annotated with four labels on a per frame basis: no compensation, lean-forward, shoulder elevation and trunk rotation. The stroke survivor patients performed two types of exercises, which were recorded with both the left and right hands: Reach-Forward-Backward and Reach-Side-to-Side. Healthy subjects completed the same scripted motions, but in addition, they simulated common compensatory movements performed by stroke survivor patients. The disadvantage of this dataset is the limited number of movements that can be performed using the rehabilitation robot.

The disadvantages of these non-image-based sensors are that they can be cumbersome for patients to wear or they require extensive resources and dedicated spaces to perform the motions. Some approaches rely on image-based sensors in order to track human motion, such as colour or depth cameras. Most of the available image-based datasets rely on a depth camera, in particular the Kinect sensor [13].

The work in [6] proposes a framework to evaluate the quality of movement recorded using a Kinect sensor. In this study, the gait of 12 healthy subjects climbing stairs was recorded along with the gait of a qualified physiotherapist simulating three scenarios of knee injury.

The dataset proposed by [7] was recorded at the Kinesiology Institute of the University of Hamburg using again a Kinect sensor. The dataset consists of 17 athletes performing three power-lifting exercises. For each routine, the athletes executed the motions both correctly and with a few typical mistakes.

The University of Idaho-Physical Rehabilitation Movement Data (UI-PRMD) dataset [5] consists of ten common physical rehabilitation exercises performed by ten healthy individuals. Each person performed ten correct and ten incorrect (nonoptimal) repetitions of the exercises. The movements were recorded using two types of sensors: a Vicon optical tracker and a Kinect sensor.

A recent collection is the KIMORE dataset reported in [14]. This dataset contains recordings of 78 subjects (44 controls and 34 patients) performing rehabilitation exercises. The collected data includes joint positions as well as RGB and depth videos. Although the dataset is a good addition to freely available resources and the authors reported how a score can be computed from the data to reflect the performance of subjects (i.e., the level of gesture correctness), the number of gestures is small and it is limited to low back pain physical exercises (the number of reported gestures is five).

The dataset presented in this article was created by recording 15 real patients with no simulated (or artificial) movements along with 14 healthy individuals, all performing repetitions of nine gestures. In comparison to our dataset, existing datasets suffer from

other limitations such as a small number of gestures or exercises restricted to specific health problems.

3. Methods

3.1. Data Acquisition

The dataset was collected using a Microsoft Kinect One sensor to record the body skeleton joints at 30 frames per second. A visual representation of the joints considered is shown in Figure 1. The dataset was acquired at Pusat Rehabilitasi Perkeso Melaka, a rehabilitation centre in Malaysia, with the help of patients and physiotherapists in the space where patients typically perform regular physiotherapy exercises. We recorded over 4.7 h of video over several days.
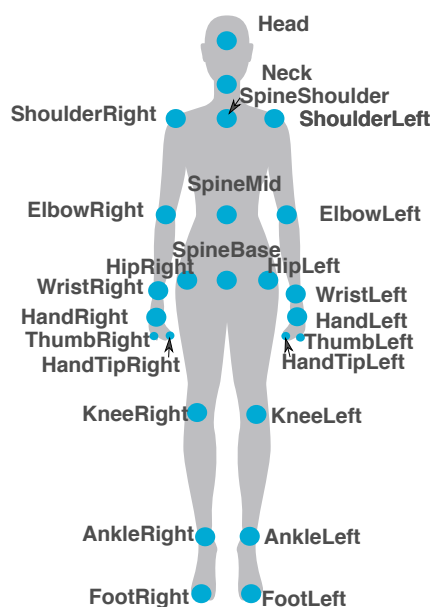

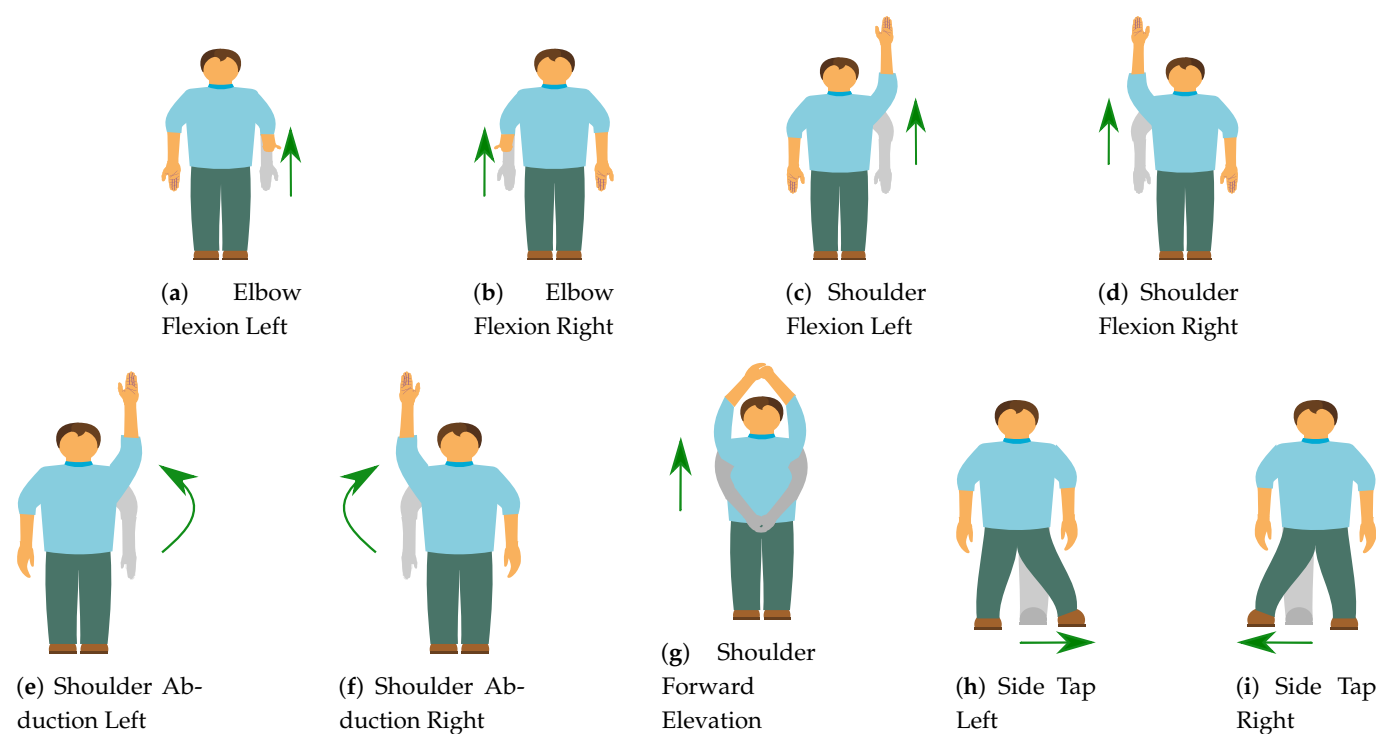**Figure 1.** Visual representation of the 3D joints extracted by Kinect.

The gestures performed by 29 subjects were captured. Out of these, 15 were patients, who were allocated IDs in the range from 201 to 216. In addition, 14 healthy individuals were recorded, out of which 7 were physiotherapists with IDs from 101 to 107 and another 7 were physiotherapy students with IDs from 301 to 307. In what follows, we refer to these 14 persons as our control group. The study was conducted ethically, conformed to the local protocol for clinical trials and obtained approval from the local ethics committee.

The patients performed the exercises in the position that was the most comfortable for them: some of them stood, while others sat in a chair or a wheelchair. To account for this variability, all of the subjects in the control group were asked to perform all of the gestures both standing and sitting in a chair.

The choice of movements was not selected for specific medical conditions but rather general simple and common movements that might be used by physiotherapists as part of a movement range assessment and rehabilitation programme. The gesture labels are represented by numbers from zero to eight, and the gesture names and brief descriptions can be found in Table 1 while a visual representation of the gestures is shown in Figure 2.

Table 1. Gesture types and description.

Gesture Index.	Gesture Name	Description
0	Elbow Flexion Left	Flexion and extension movement of the left elbow joint
1	Elbow Flexion Right	Flexion and extension movement of the right elbow joint
2	Shoulder Flexion Left	Flexion and extension movement of left shoulder while keeping the arm straight in front of the body
3	Shoulder Flexion Right	Flexion and extension movement of right shoulder while keeping the arm straight in front of the body
4	Shoulder Abduction Left	The left arm is raised away from the side of the body while keeping the arm straight
5	Shoulder Abduction Right	The right arm is raised away from the side of the body while keeping the arm straight
6	Shoulder Forward Elevation	With hands clap together, the arms are kept straight and raised above the head, keeping the elbows straight
7	Side tap Left	The left leg is moved to the left side and back while keeping the balance
8	Side tap Right	The right leg is moved to the right side and back while maintaining balance

(**a**) Elbow Flexion Left

(**b**) Elbow Flexion Right

(**c**) Shoulder Flexion Left

(**d**) Shoulder Flexion Right

(**e**) Shoulder Abduction Left

(**f**) Shoulder Abduction Right

(**g**) Shoulder Forward Elevation

(**h**) Side Tap Left

(**i**) Side Tap Right

Figure 2. Examples of the nine movements.

3.2. Information about the Subjects

Table 2 contains the demographic information about the 15 patients we recorded, while Table 3 contains information about the healthy subjects. The average age for the patients is 43 years, while the average age for the healthy subjects is approximately 26 years. The health condition and the diagnostic of the patients is diverse, with different parts of the body being affected. The wheelchair column only refers to the fact that the patient used or did not use a wheelchair during the data collection stage and does not represent a permanent condition. Five of the patients suffered a spinal cord injury, five of them suffered strokes, one of them suffered a brain injury, another one had a neurological condition, one

suffered from arm injury, one had a fractured femur and one had a knee-level amputation (the patient wore a prosthetic leg).

Table 2. Patient information.

Patient ID	Age-Group	Sex	Body Side Affected	Wheelchair
201	50–59	male	Lower body	Yes
202	50–59	male	Left	No
203	20–29	male	Lower body	Yes
204	60+	male	n/a	No
205	30–39	female	Left	No
206	30–39	female	Right	No
207	40–49	male	Right	No
209	20–29	male	Right	Yes
210	60+	male	Upper and lower limb weakness	No
211	50–59	female	Left	No
212	40–49	male	Upper and lower limb weakness	Yes
213	50–59	male	Right	No
214	20–29	male	n/a	No
215	30–39	female	Right	Sometimes
216	30–39	male	n/a	Yes

Table 3. Healthy persons information.

Person ID	Age-Group	Sex
101	30–39	male
102	30–39	female
103	20–29	male
104	30–39	female
105	20–29	female
106	20–29	male
107	20–29	male
301	20–29	male
302	20–29	male
303	20–29	female
304	20–29	female
305	20–29	male
306	20–29	female
307	20–29	female

4. Data Records

The dataset released contains 2589 files, with each file corresponding to one gesture. The nomenclature of the files is as follows:

SubjectID_DateID_GestureLabel_RepetitionNo_CorrectLabel_Position.txt

For example, the file 303_18_4_10_1_*stand*.txt refers to the gesture performed by the person with ID 303, on the date labelled with ID 18, on the 10th repetition of the gesture labelled 4, and performed correctly while standing. Each file has an associated *CorrectLabel* that can have the values 1, for a correct gesture, 2, for an incorrect gesture, and 3 for gestures that are incorrect but poorly executed and, based only on the recording, would be impossible to assign a gesture label. For the analysis that follows, we ignore the files with *CorrectLabel* 3 (there are only 12 files with this label); however, because all of these movements were performed by patients, they might be useful for certain types of movement modelling and transfer learning, so we left them in the final dataset. The rest of the analysis in this article refers only to the 2577 files with correctness labels 1 and 2. It is

worth mentioning that the correctness labelling is binary (the gesture is either correct or incorrect) and not discrete (measuring the level of correctness).

Out of a total of 2577 gesture sequences, 1215 were performed standing, 952 were performed sitting on a chair, 359 were performed sitting on a wheelchair, and 51 were performed using a stand frame for support.

We provide the data in two formats. The first one is a simplified comma-separated value format with each line containing the 3D coordinates of the 25 joints. The second format is a raw data file where, in addition to the 3D coordinates, we include a timestamp for every frame, information for every joint mentioning whether the joint is tracked, and the 2D projections of the 3D coordinates.

The data contents can be described as follows: (i) each clip contains n frames, (ii) each frame contains spatial information of m joints (in our case 25), and (iii) each joint is represented by three axes (x, y, z). Hence, the total number of features is 75.

Along with the 3D coordinates of the 25 joints, we provide also the raw depth map images with the same nomenclature as the corresponding .csv file.

Data Variations

As the data were collected from real patients, a significant degree of variability is expected. We refer to the variability within the same move repeated by the same subject multiple times as the *within variability*. In addition, we refer to the variability between different subjects repeating a particular move as the *between variability*.

An example of the *within variability* is shown in Figure 3, where the x-axis of the right wrist of subject 103 (a physiotherapist) performing gesture 5 (right shoulder abduction) correctly while standing is plotted (please notice that the data were normalised by subtracting the spine-base's x-axis). As it can be seen, the data vary not only in length (i.e., the number of frames) but also in position (coordinates) values.
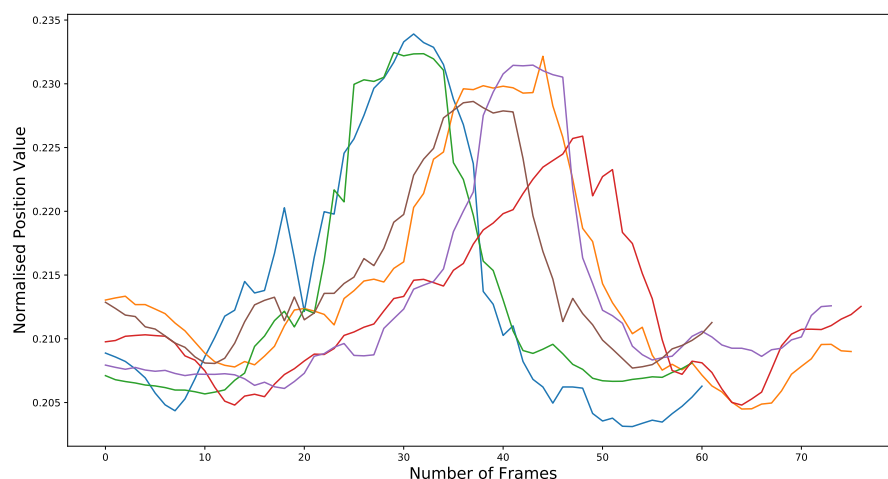

Figure 3. Example of right wrist variations in gesture number 5 performed correctly by the same person (6 repetitions).

Five physiotherapists performed gesture 5 correctly several times. In order to examine their variability, we normalised all of their data for this move to the same number of frames (i.e., 100 frames) using cubic interpolation. We then averaged the x-axis values for each repetition per subject (after subtracting the spine-base's x-axis) and plotted the results in Figure 4. As can be seen, it is obvious from the figure that, indeed, there is a large degree in variation between subjects. Nevertheless, there is an overall trend in how the movement is performed: the right wrist starts from a low position, moves upwards, and returns to the original position.
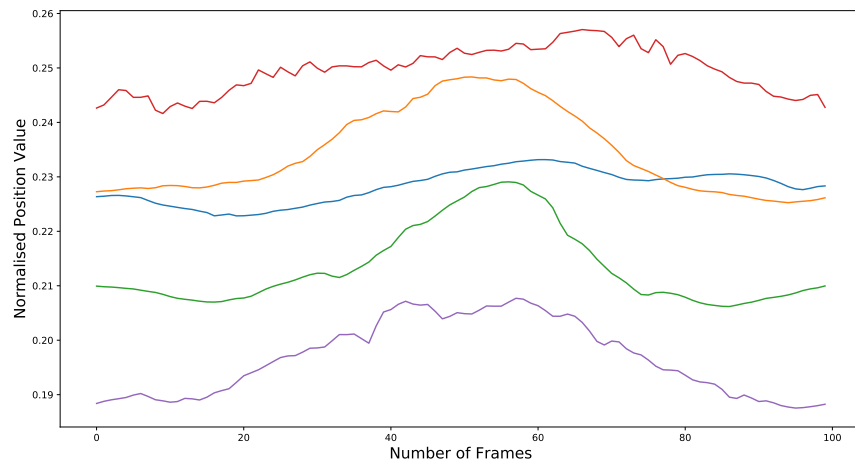
Figure 4. Example of right wrist variations in gesture number 5 performed correctly by 5 different persons (average of repetitions).

5. Data Distribution and Augmentation

As mentioned in Section 3.1, we recorded 14 healthy individuals as our controls (most of whom are physiotherapists) performing the same gestures. Because patients have various physical limitations, not all of them completed the same number of gesture repetitions (i.e., episodes). The same applies for controls as they were not all available for the same amount of time. Each subject attempted to perform gestures a number of times. It is these repetitions that are labelled as correct or incorrect. The number of correct and incorrect repetitions for each gesture is shown in Figure 5.
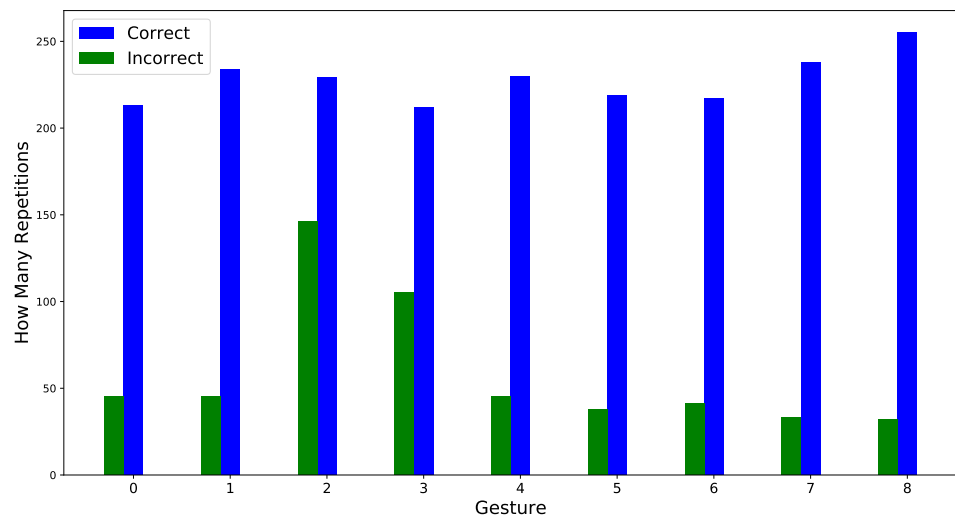
Figure 5. Number of repetitions for each gesture.

In these recordings, the correct repetitions were mostly performed by the controls, although many patients were able to perform some of the repetitions correctly. Therefore, the distribution of correct vs. incorrect repetitions can differ from one gesture to another, as shown in Figure 6.
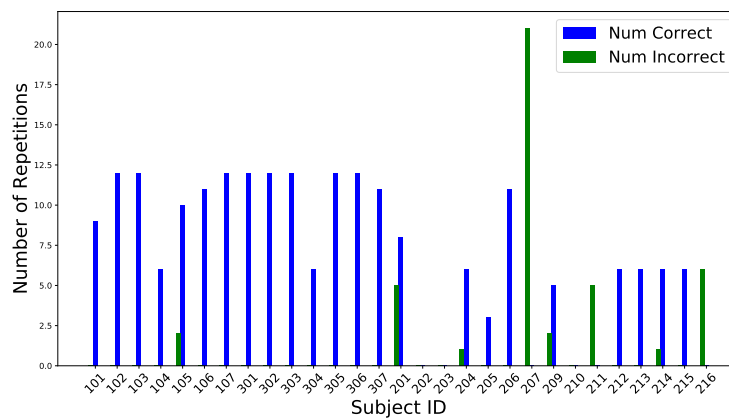
(**a**) Gesture 0: elbow flexion left
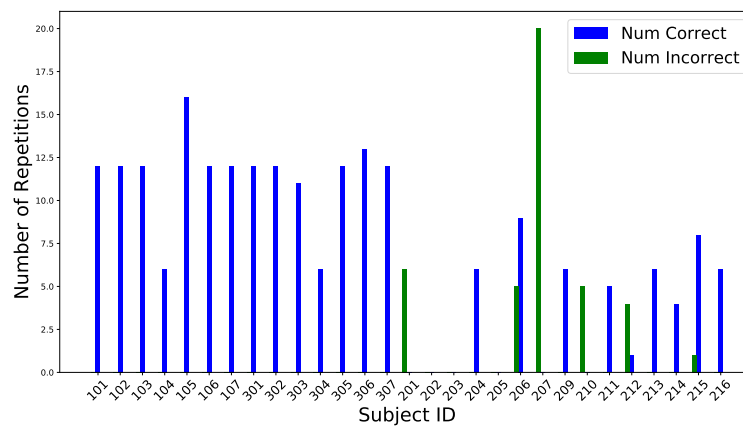
(**b**) Gesture 6: frontal both hands

Figure 6. Distribution of repetition counts per subjects in two gestures.

The data are highly unbalanced. That is, the distribution of different classes and categories is different (e.g., the number of correct and incorrect moves is unequal). The distribution of repetitions for each gesture is shown in Figure 5. As it can be seen, there are far more correct moves than incorrect ones. Hence, to balance the data, either some correct moves can be removed or more incorrect moves can be recorded. The first option means that we lose data, and therefore, it should be avoided. The second option is costly as it is not always easy to find real patients who are willing to perform movements and to be recorded. Based on this, a third option would be to generate synthetic data that belongs to the incorrect moves (i.e., data that have similar characteristics to the incorrect move data).

A number of time-series data augmentation techniques is reported in the literature. For example, various architectures of generative adversarial networks (GANs) were used in [15] to augment and classify gesture data as correct or incorrect. Another set of techniques is provided in [16]. These techniques are based on geometric and affine transformations such as rotation and time warping. They also include simple methods such as adding random noise, scaling, and jittering. Please observe that, because the code to generate new augmentation data is freely available and easy to use, we do not provide any augmentation data. Another reason is that each time the code is run, slightly different data are generated.

6. Technical Validation

In the proposed dataset, the minimum length of a gesture sequence (measured as the number of frames) is 13, while the maximum length is 1586. On average, a gesture has 84 frames and 75% of the data has a length below 89 frames. There is a strong tendency for

incorrect gestures (on average, 148 frames) to be longer than the correct ones (on average, 68 frames; see Figure 7).

Using the sensor recording speed of 30 frames per second, on average, the minimum length of a gesture is 0.3 s and the maximum one is 52 s. The length of the correct gestures is no longer than 13 s, while a total number of 25 incorrect gestures (4.7% of total incorrect gestures) have a length longer than this value. This is most likely due to either the patient struggling to perform the gesture or taking a long time to prepare for the gesture. Although these situations can be considered outliers, we decided to keep these recordings in the dataset.

As seen in Figure 8, most of the incorrect gestures have a duration significantly longer than the correct executions, with gestures 2 and 3 being the most obvious ones.
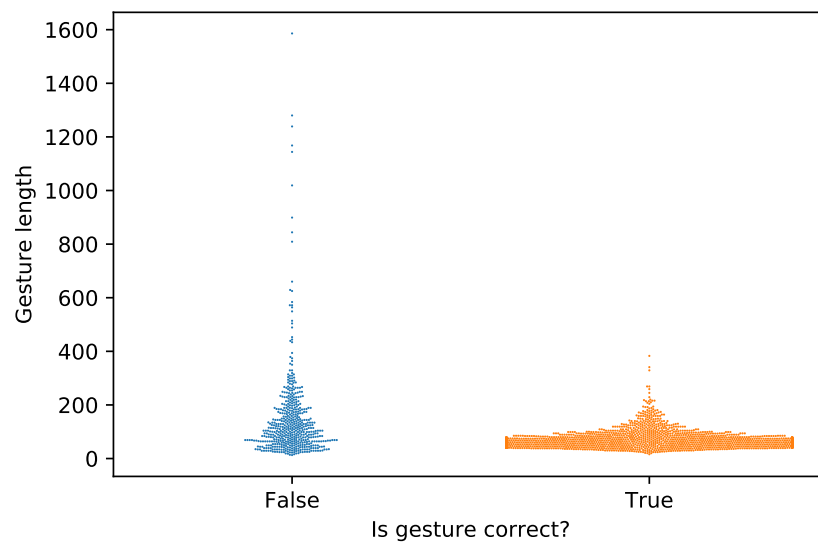
Figure 7. Visualisation of gesture length with respect to gesture correctness label.
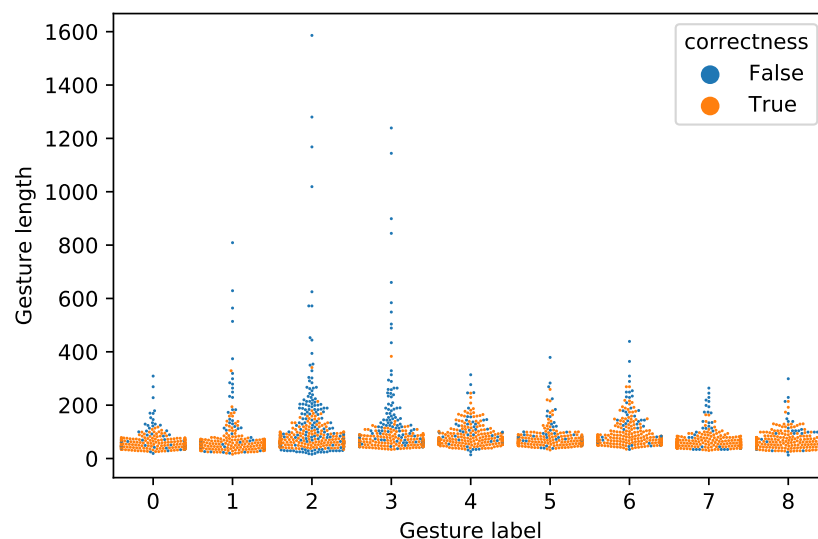
Figure 8. Visualisation of gesture length with respect to gesture correctness label and gesture type.

Each healthy subject repeated most of the gestures at least five times. In what concerns the patients, some of them were not able to perform some of the gestures. For example, the subject with ID 205 could not perform shoulder forward elevation due to a left arm injury. Overall, the patients repeated the gestures to the best of their ability. Figure 9 displays an overall visualisation of the number of repetitions for each gesture by each subject. As it can be observed, some patients repeated the exercises for much longer than

they were instructed or wanted to come back for several recording sessions. In Figure 10, the distribution of the incorrect execution of different gestures is presented. As it can be expected, the majority of the incorrect gestures (98% of them) were performed by patients while the control group had very few incorrect gesture executions.
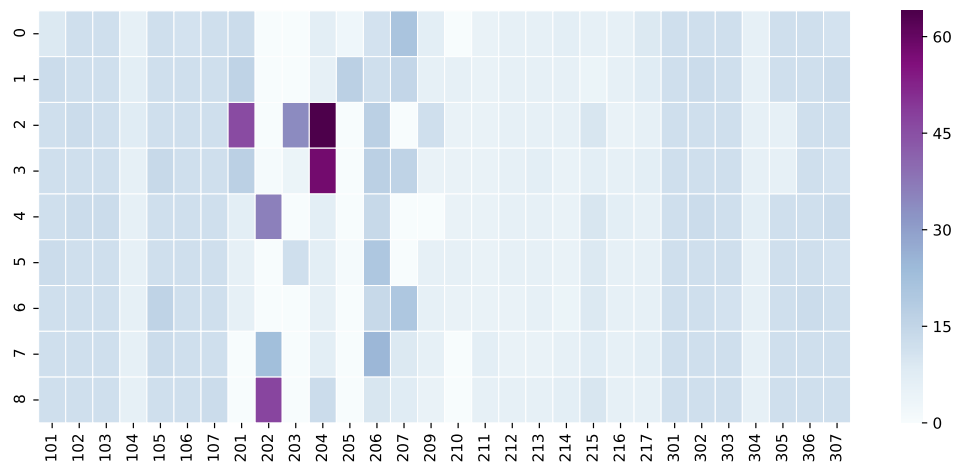

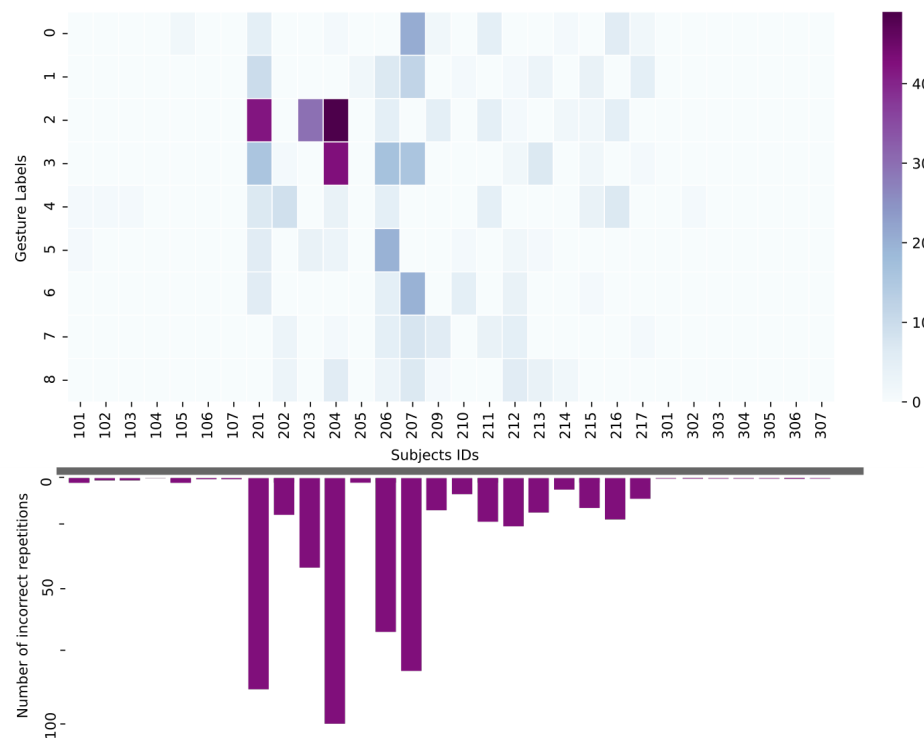**Figure 9.** Heat map of number of gestures performed by every subject.

Figure 10. Heat map of number of incorrect gestures performed by every subject and overall counts of incorrect movements per subject.

Correctness, especially when referring to how well a gesture is performed, can be a highly subjective measure. Two annotators reviewed each recording and independently annotated each gesture as being correct or not. The inner-annotator agreement is 88%. In total, 290 recordings were revisited by both annotators, and a final decision was made regarding the correctness label.

In Figure 11, we present a few examples of correct and incorrect executions for the shoulder flexion left exercise. A correct execution, in this case, involves flexion and extension of the left shoulder while keeping the arm straight in front of the body. The arm

should be raised straight above the head. An incorrect execution is considered when the elbow is bent, the arm is not raised high enough or the movement was compensating by swinging the arm. In Figure 11, we show an overlaid skeleton representation in time of the recorded 3D joint points for an individual gesture repetition. To represent movement, the skeleton is drawn using shades of green for up to half of the movement and shades of red for the second half of the movement.

Figure 11. Examples of the correct (first row) and incorrect (second row) left shoulder flexion executions.

7. Discussion and Conclusions

The contribution of this paper is the presentation of a dataset of movements related to nine physical rehabilitation exercises. The gestures were performed by 29 subjects, out of which 15 patients and 14 healthy control were annotated by gesture type, position and a correctness label. As with all datasets, there are some limitations. The gestures are not associated with a particular condition, with the patients experiencing a variety of conditions, from stroke to spinal cord injury. Although we strove to collect as much data as possible, we only collected data from 15 patients. This is still larger than other existing datasets such as [5], where 10 healthy people were recorded, and [12], where they had 10 stroke patients, but the size of the dataset may be a shortcoming in the context of using machine learning methods. Another possible limitation is the discontinuity of the Kinect sensor, although other similar depth cameras are still available (Intel Depth Cameras [17] and Orbec Astra [18]). In the context of limited availability of gesture-related datasets that contain real patient movements, we envision this dataset to be used either on its own or in combination with other datasets, especially with the rapid expansion of the field of transfer learning [19].

References

1. Kinect—Windows App Development. 2019. Available online: https://developer.microsoft.com/en-us/windows/kinect (accessed on 29 April 2021).
2. Kitsunezaki, N.; Adachi, E.; Masuda, T.; Mizusawa, J. KINECT applications for the physical rehabilitation. In Proceedings of the 2013 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Gatineau, QC, Canada, 4–5 May 2013; pp. 294–299.
3. Su, B.; Wu, H.; Sheng, M.; Shen, C. Accurate Hierarchical Human Actions Recognition From Kinect Skeleton Data. *IEEE Access* **2019**, *7*, 52532–52541.
4. Chiuchisan, I.; Geman, O.; Postolache, O. Future trends in exergaming using MS Kinect for medical rehabilitation. In Proceedings of the 2018 International Conference and Exposition on Electrical And Power Engineering (EPE), Iasi, Romania, 18–19 October 2018; pp. 0683–0687.
5. Vakanski, A.; Jun, H.P.; Paul, D.; Baker, R. A data set of human body movements for physical rehabilitation exercises. *Data* **2018**, *3*, 2.
6. Paiement, A.; Tao, L.; Hannuna, S.; Camplani, M.; Damen, D.; Mirmehdi, M. Online quality assessment of human movement from skeleton data. In Proceedings of the British Machine Vision Conference, Nottingham, UK, 1–5 September 2014; pp. 153–166.
7. Parisi, G.I.; von Stosch, F.; Magg, S.; Wermter, S. Learning human motion feedback with neural self-organization. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–16 July 2015; pp. 1–6.
8. Devineau, G.; Moutarde, F.; Xi, W.; Yang, J. Deep learning for hand gesture recognition on skeletal data. In Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 15–19 May 2018; pp. 106–113.
9. Niewiadomski, R.; Kolykhalova, K.; Piana, S.; Alborno, P.; Volpe, G.; Camurri, A. Analysis of movement quality in full-body physical activities. *ACM Trans. Interact. Intell. Syst. TiiS* **2019**, *9*, 1.
10. Hussain, R.G.; Ghazanfar, M.A.; Azam, M.A.; Naeem, U.; Rehman, S.U. A performance comparison of machine learning classification approaches for robust activity of daily living recognition. *Artif. Intell. Rev.* **2019**, *52*, 357–379.
11. Ebert, A.; Beck, M.T.; Mattausch, A.; Belzner, L.; Linnhoff-Popien, C. Qualitative assessment of recurrent human motion. In Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO), Kos Island, Greece, 28 August–2 September 2017; pp. 306–310.
12. Dolatabadi, E.; Zhi, Y.X.; Ye, B.; Coahran, M.; Lupinacci, G.; Mihailidis, A.; Taati, B. The toronto rehab stroke pose dataset to detect compensation during stroke rehabilitation therapy. In Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare, Barcelona, Spain, 23–26 May 2017; pp. 375–381.
13. Lun, R.; Zhao, W. A survey of applications and human motion recognition with microsoft kinect. *Int. J. Pattern Recognit. Artif. Intell.* **2015**, *29*, 1555008.

14. Capecci, M.; Ceravolo, M.G.; Ferracuti, F.; Iarlori, S.; Monteriù, A.; Romeo, L.; Verdini, F. The KIMORE Dataset: KInematic Assessment of MOvement and Clinical Scores for Remote Monitoring of Physical REhabilitation. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2019, 27, 1436–1448.
15. Li, L.; Vakanski, A. Generative Adversarial Networks for Generation and Classification of Physical Rehabilitation Movement Episodes. *Int. J. Mach. Learn. Comput.* **2018**, *8*, 428–436.
16. Um, T.T.; Pfister, F.M.; Pichler, D.; Endo, S.; Lang, M.; Hirche, S.; Kulić, D. Data Augmentation of Wearable Sensor Data for Parkinson's Disease Monitoring Using Convolutional Neural Networks. In Proceedings of the 19th ACM International Conference on Multimodal Interaction, Glasgow, UK, 13–17 November 2017; pp. 216–220.
17. Intel® RealSense™ Depth and Tracking Cameras. 2021. Available online: https://www.intelrealsense.com/depth-camera-d435/ (accessed on 22 April 2021).
18. Orbec Astra. 2021. Available online: https://orbbec3d.com (accessed on 22 April 2021).
19. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; He, Q. A comprehensive survey on transfer learning. *Proc. IEEE* **2020**, *109*, 43–76.