The KIMORE Dataset: KInematic Assessment of MOvement and Clinical Scores for Remote Monitoring of Physical REhabilitation

Marianna Capecci, Maria Gabriella Ceravolo, Francesco Ferracuti, *Member, IEEE*, Sabrina Iarlori, Andrea Monteriù, *Member, IEEE*, Luca Romeo, and Federica Verdini

***Abstract*— This paper proposes a free dataset, available at the following link,[1] named KIMORE, regarding different rehabilitation exercises collected by a RGB-D sensor. Three data inputs including RGB, depth videos, and skeleton joint positions were recorded during five physical exercises, specific for low back pain and accurately selected by physicians. For each exercise, the dataset also provides a set of features, specifically defined by the physicians, and relevant to describe its scope. These features, validated with respect to a stereophotogrammetric system, can be analyzed to compute a score for the subject's performance. The dataset also contains an evaluation of the same performance provided by the clinicians, through a clinical questionnaire. The impact of KIMORE has been analyzed by comparing the output obtained by an example of rule and template-based approaches and the clinical score. The dataset presented is intended to be used as a benchmark for human movement assessment in a rehabilitation scenario in order to test the effectiveness and the reliability of different computational approaches. Unlike other existing datasets, the KIMORE merges a large heterogeneous population of 78 subjects, divided into 2 groups with 44 healthy subjects and 34 with motor dysfunctions. It provides the most clinically-relevant features and the clinical score for each exercise.**

***Index Terms*— Dataset, rehabilitation, motion analysis, RGB-D sensor.**

I. Introduction

In recent years, chronic diseases have been affecting the quality of life of many people, leading to progressive limitations in motor activities and reducing participation in social life when a correct and continuous rehabilitation support is not provided. In this context, an accurate and reliable rehabilitation framework is necessary to reduce the high demands for healthcare staff as well as to make rehabilitation more enjoyable and acceptable in terms of adherence, monitoring, access and sustainability [1]. New modalities of health service delivery have proliferated, providing remarkable solutions for overcoming related issues and offering individualized programs beyond the hospital setting, such as telerehabilitation. This approach is based on regular monitoring of the patient's state of health and progress, with respect to the aim of the treatment and to his/her expectations [2]–[5]. Generally, a rehabilitation program is delivered in clinical facilities or at home by a physiotherapist, who continuously provides feedback on gesture accuracy, in terms of goal, motion and posture, in order to obtain the best result as regards safety and efficacy in the short as well as in the long term. The feedback/information about gesture accuracy while performing or after a movement promotes motor learning and retention, minimizing possible side effects and maximizing physical benefits [6]. Therefore, an effective and safe telerehabilitation architecture should guarantee the same supervision of the gesture in order to reach results similar to those provided by the therapist. In this context, both postural alignment and kinematics should be monitored according to [7]. Different telerehabilitation approaches require wearable and/or vision-based systems to monitor the patient during the exercise. Particularly, Red-Green-Blue Depth (RGB-D) cameras can be used as low-cost markerless systems to analyze human motion and support physiotherapists in the rehabilitation cycle [8]–[16].

In this context, the motion capture system technology, adopted to record the subject's performance, can produce reliable results, if complemented by specific and accurate data processing algorithms. Machine learning algorithms have often been employed to perform motion assessment. Some of these try to provide movement evaluation in order to give a feedback about the correct execution of the gesture performance. In literature, it is possible to find the two most important human motion assessment approaches [17]: *rule-* and *template*-based. While in the *rule*-based approach [17]–[20], experts (e.g., medical staff) identify some motion descriptors (e.g., angles,

Manuscript received December 5, 2018; revised March 27, 2019 and June 11, 2019; accepted June 11, 2019. Date of publication June 14, 2019; date of current version July 4, 2019. *(Corresponding author: Luca Romeo.)*

M. Capecci and M. G. Ceravolo are with the Department of Clinical and Experimental Medicine, Università Politecnica delle Marche, 60131 Ancona, Italy (e-mail: m.capecci@univpm.it; m.g.ceravolo@univpm.it).

F. Ferracuti, S. Iarlori, A. Monteriù, and F. Verdini are with the Department of Information Engineering, Università Politecnica delle Marche, 60131 Ancona, Italy (e-mail: f.ferracuti@univpm.it; s.iarlori@univpm.it; a.monteriu@univpm.it; f.verdini@univpm.it).

L. Romeo is with the Department of Information Engineering, Università Politecnica delle Marche, 60131 Ancona, Italy, and also with the Cognition, Motion and Neuroscience and Computational Statistics and Machine Learning, Fondazione Istituto Italiano di Tecnologia, 16163 Genova, Italy (e-mail: l.romeo@univpm.it).

Digital Object Identifier 10.1109/TNSRE.2019.2923060

[1]https://univpm-my.sharepoint.com/:f:/g/personal/p008099_staff_univpm_it/EiwbKIzk6N9NoJQx4J8aubIBx0o7tIa1XwclWp1NmRkA-w?e=F3jtBk

joint position, relative distance, velocity), which define the "motion sample", in the *template*-based approach [12], [13], [21]–[29], the motion sequence is recorded *a priori*, and then used as an exemplar to be compared with the observations, through action similarity approaches or using machine learning methodologies, which allow it to be easily generalized to different types of exercise. Machine learning based methodologies require a huge amount of data in order to learn the correct motion sequence as well as all the possible errors. The algorithm should be able to generalize across different subjects, pathologies or sets of exercises. Accordingly, template-based approaches require a completely supervised setting, where the training data are composed of examples of the input vectors along with their corresponding output (i.e., score for the performance of the exercise). To the best of the authors' knowledge, very few datasets [30] overcome the costs associated with the labeling process, providing this target information annotated by clinicians. In this scenario, the current paper contributes to the creation of a freely available dataset, named KIMORE (KInematic assessment of MOvement for remote monitoring of physical REhabilitation), composed of rehabilitation exercises, collected by the Microsoft Kinect v2 sensor, with both healthy and disabled subjects. The dataset consists of three synchronized data typologies, which include RGB, Depth videos, skeleton positions and orientations in the format produced by the skeleton tracking system. Skeleton data (i.e., trajectory position and orientation of virtual joints), are computed through video streams, using the algorithm proposed by [31]. The validation of these virtual joint angles has been the topic of several studies [32], [33] which highlight an high level of agreement between the Kinect-based motion capture system (i.e., skeleton tracking algorithm) and the ground truth system (i.e., stereophotogrammetric system) when tracking joints displacement. More importantly, the findings behind these studies suggest that the Kinect-based motion capture system may be a viable alternative to professional three-dimensional systems for certain applications [33].

Two physicians, specialized in Physical and Rehabilitation Medicine, selected five exercises usually adopted in rehabilitation programs for low back pain [34], [35], for a study aimed at carrying out exercise assessment in a visual markerless scenario [36], [37]. The dataset can be categorized according to the definition introduced in [38], as a single view action/activity, where each action, performed by one actor at a time, is captured from a single specific viewpoint, to distinguish it from multi-view action/activity datasets [38] and from human-human interaction/multi-person activity datasets [38]. Together with the raw data, clinically relevant motion features, suggested by physicians and validated with respect to a stereophotogrammetric system [32], [37], are provided in the dataset. These features and the relative identified trajectories can be used to study the task and subsequently to test and compare rule- and template- based approaches for physical exercises assessment [36], [37]. Moreover, the clinical evaluation, based on a standard questionnaire designed by physicians, is reported with the references for the validation. KIMORE provides a score which is useful for (i) properly designing the template-based algorithm for movement assessment and (ii) measuring the performance of the proposed approaches with respect to the clinical ground truth evaluation. Taking into account this aspect and considering the recent interruption of Kinect on the market, our work aims to provide a dataset which:

- includes RGB, depth and skeleton joints positions and orientations. The dataset also comprises clinical features, which are invariant among people and selected on the basis of the scope of the exercise (see [37], for more detail [18]) as well as the Matlab code to compute it;
- can be easily generalizable to different domains (i.e., no visual tracking sensors) which involve not only the validated clinical features. In particular, although Kinect produces a stream of body segment orientations, these measurements must be numerically manipulated to yield clinically relevant kinematic data [33].

The paper is organized as follows. A comparison with other studies and datasets is provided in subsection I-A. The population involved, the exercises description, the relative feature extraction, the clinical assessment, with a description of the questionnaire and how the data are organized in the dataset are introduced in Section II. Section III presents the results related to the statistical analysis of the scores obtained, through both the clinical and the Kinect-based assessment. Finally, the discussion and conclusions are presented in Sections IV and V.

A. Related Works

Many data-driven approaches are based on relatively large training data. Thus, recently there has been an increased interest in collecting datasets in various application domains ranging from brain computer interfaces [39]–[41] to dietary monitoring and food recognition [42] to predictive telediagnosis and telemonitoring [43]. The existing datasets, related to human motion data were created using RGB-D or similar sensors as in [38] and [44] to facilitate the development and the approach of new algorithms, using accurate depth information or simply the RGB data. Generally, the data have been compared with those collected by other kinds of sensors like the inertial sensor in [45] to test multimodal approaches and develop multimodal datasets [46], [47]. In particular, the diffusion of the Kinect sensor and its accessibility, has allowed the creation of different datasets on the basis of gesture, primitive movements, action and activities as in [48] and [49]. This sensor has also been used as a camera for monitoring emotions related to pain in a rehabilitation context [50]. The datasets, presented in [51] and [52], include different activities: in the MADS dataset, complex poses, related to Martial Arts, Dancing and Sports are available to test motion tracking algorithms. In [53] and [54], the authors propose two datasets, the ReadingAct and the RGBD-HuDaAct, for human activity recognition using the Kinect sensor. In both papers, feature extraction is performed considering the kinematics of the movement and not the clinical aspects of the gesture (task- vs clinical-oriented), defined only on the basis of the scope of the movement and not with respect to clinically relevant dynamics. The applications of RGB-D-based action

TABLE I

COMPARISON BETWEEN KIMORE AND THE DATASETS PRESENTED IN LITERATURE, IN THE REHABILITATION CONTEXT. THE MAIN PHASES OF DATA COLLECTION, FEATURE EXTRACTION, PREPROCESSING, DATA VALIDATION AND EXERCISE EVALUATION, ARE INCLUDED AND CONTRASTED ACROSS THE DIFFERENT DATASETS. FURTHERMORE, THE TABLE REPORTS WHETHER THE DATASET IS PUBLISHED AND AVAILABLE FREE OF CHARGE. THE INFORMATION ABOUT THE ENROLLED POPULATION AND THE TYPE OF DATA ACQUIRED, RGB, DEPTH (D), JOINT POSITIONS (JP) AND JOINT ORIENTATIONS (JO), COMPLETES THE DESCRIPTION OF THE PROPOSED DATASET

Dataset References	Workout SU-10 [56]	K3Da [57]	UI-PRMD [55]	Multimodal Dataset [58]	KiReS Dataset [30]	**KIMORE**
Available	✓	✓	✓	✓		✓
Data (RGB, D, JP, JO)	D, JP, JO	D, JP, JO	JP, JO	JP, JO		✓
Population (subjects)	15	54	10	21	5	78
Published	✓	✓	✓	✓		✓
Validation			✓			✓
Preprocessing			✓			✓
Exercises	✓		✓	✓	✓	✓
Features Extraction	✓	✓			✓	✓
Exercise Evaluation	✓	✓	✓	✓	✓	✓
Clinical Evaluation						✓

datasets, in the literature, are limited for different reasons: the type and dimension of the analyzed population, the validation and preprocessing (see Table I), and the restricted types of exercises and motor tasks included in each dataset. These characteristics may reduce the dataset applicability [38].

In the present work, the authors propose a comparison of the KIMORE dataset with some of those available in literature, focused on the rehabilitation context. In detail, Table I summarizes the characteristics of the proposed dataset with respect to the others cited [30], [55]–[58] that compile rehabilitation exercises, with the exception of the dataset in [57] that presents the movement assessment of standardized tests selected by clinicians. The authors identified the most salient factors, described in Table I and chosen according to those discussed in [38] (i.e., dataset size, applicability, evaluation protocols) in order to measure and compare the reliability of the proposed dataset with respect to other works. All the steps in Table I were previously carried out with the proposed KIMORE dataset in [18], [32], [36], and [37]s. Differently from the other datasets, KIMORE provides the two main data modalities (i.e., RGB and Depth). This aspect opens up a whole range of possibilities for testing several computer vision approaches which are not directly based on clinical features and Kinect skeleton tracking. For instance, open frameworks, such as OpenPose [59], may be used to obtain virtual skeleton joints directly from RGB data.

With respect to the enrolled population, the analyzed datasets, coming from the literature, included smaller samples, ranging from a minimum of 5 subjects, as in [30], to a maximum of 54 in [57] with a prevalence of healthy subjects and a limited age range; conversely, in the proposed dataset, the number of people involved in the study, 78, is larger than in other works [30], [55]–[58] and the subjects display a wider range of age and health/disability conditions.

Note that, almost all the reported datasets are accessible and published, except for the one proposed in [30], which was developed with the sole purpose to allow physiotherapists and patients to test the prototype of the telerehabilitation system. It is worth bearing in mind that the proposed dataset has been validated in [32] and that the preprocessing step described here is not reported in the presentation of other available works; they do not present data or specific information related to preprocessing.

Although the aim of many papers showing a dataset is to obtain an evaluation of the subject's movement, only a few of them present a feature extraction method [30], [56], [57] to provide performance assessment. Differently from the approach introduced in this study, Negin *et al.* [56] and Leightley *et al.* [57] proposed a feature selection based on a decision forest [56] and a k-means clustering [57]. Feature extraction, encapsulating prior clinical knowledge related to the objective and kinematic constraints of physical exercises, is chosen by these authors to obtain salient motion features for movement assessment, while all the works introduce a direct exercise evaluation through the different machine learning methods adopted (i.e., support vector machine, artificial neural networks).

In the cited survey [52], the authors observed that size, applicability, feasibility of ground truth labels and evaluation protocols are lacking in the available literature on RGB-D-based datasets, notwithstanding the importance of providing a reliable tool for motion analysis supporting rehabilitation, as emerged from preliminary reports [60]. They showed that detection supported exercise therapy produced similar or even better enhanced clinical outcomes compared to conventional exercise therapy [61]. From this perspective, only the proposed KIMORE dataset includes annotations made by expert clinicians, of the same exercises performed by the different enrolled subjects,through the compilation of a designed questionnaire which is reported in Section II-F and published in [18]. Moreover, the assessment is not only related to all the exercises, but the total score is obtained by averaging the local scores, related to the primary outcome, and the kinematic constraints described for each exercise. The contribution of a medical staff, to describe the features and to evaluate the performance, is not introduced in any other work.

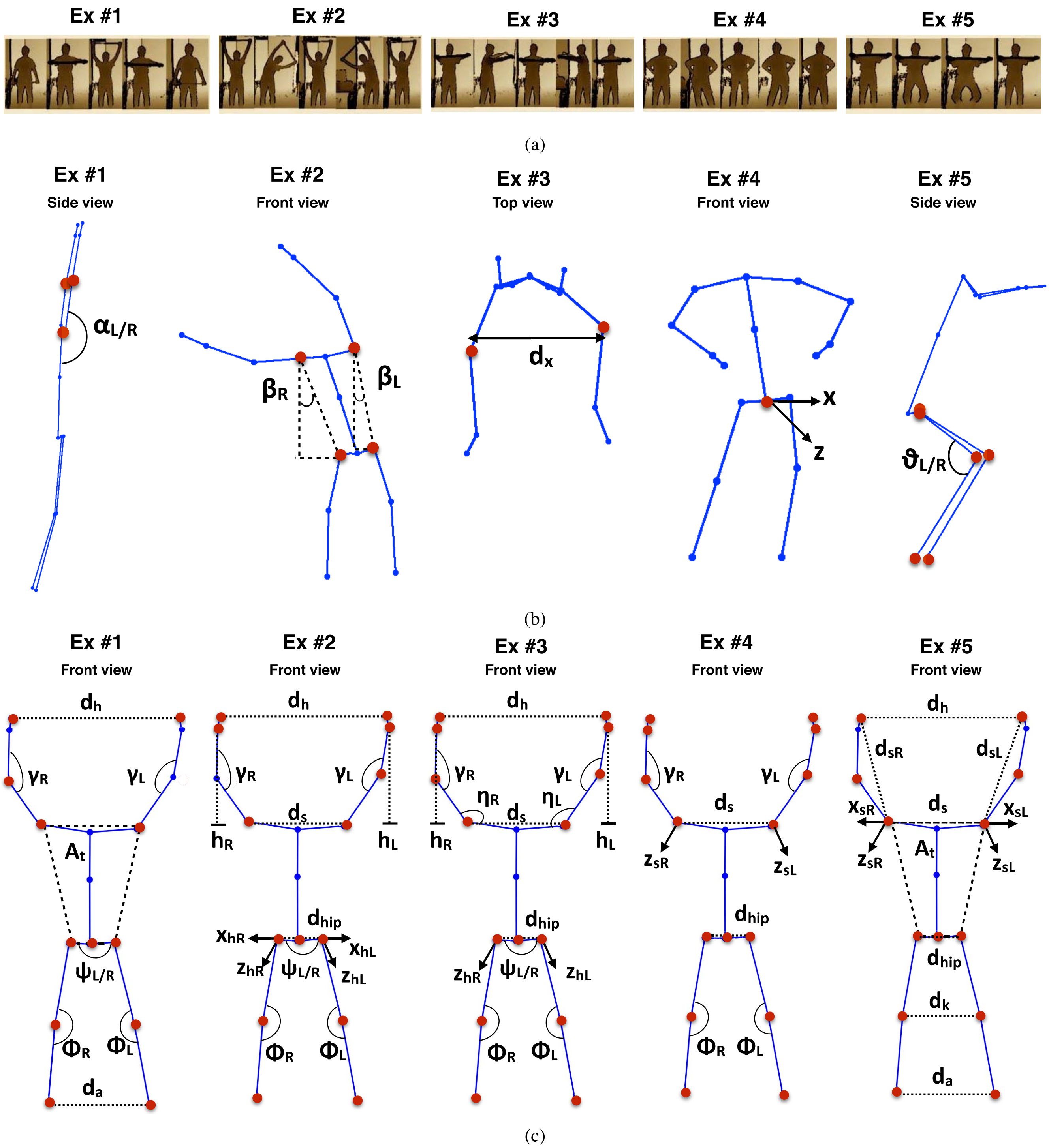


Fig. 1. The 5 rehabilitation Exercises: PO and CF features extraction. (a) Exercises. (b) POs extracted. (c) CFs extracted.

II. Materials and Methods

A. Population

The authors enrolled 44 healthy subjects, with no history of neurological or musculoskeletal problems and no recent traumas. The average age was 35 years (mean (SD) = 36.7 (16.8) years) and 15 subjects were females. The healthy subjects contributed to defining the normative data of the dataset and constituted the *Control Group (CG)*. Within this group, 12 subjects were physiotherapists and experts in the rehabilitation of back pain and postural disorders *CG-E*, while the remaining 32 were non-expert healthy subjects *CG-NE*.

In addition to the healthy group, 34 other subjects were also enrolled. They were suffering from chronic motor disabilities due to different pathologies affecting posture and causing back pain (stroke (n= 10), Parkinson's disease (n= 16), back pain due to spondylosis (n = 8)). The whole *Group* with *Pain and Postural disorders (GPP)* was 60 years old on average (mean (SD)= 60.44 (14.2) years) and was constituted by 19 women and 15 men. No subject in the GPP was in an acute phase of illness; patients were consecutively enrolled at a neurorehabilitation facility during follow-up visits.

All the enrolled subjects performed the experimental protocol and signed informed consent for data publication. Among them, 8 subjects of CG and 15 of GPP did not allow the publication of either RGB or depth videos while the related kinematic data are present in the KIMORE dataset.

The study conformed to the Helsinki protocol for clinical trials and was approved by the local ethics committee at the University Hospital in Ancona.

B. Data Collection: Kinect v2

A RGB-D camera allows a 3D structure of the scene to be computed, with good invariance against illumination changes, color and texture. In order to collect the dataset, a Microsoft Kinect v2 was adopted as a RGB-D vision sensor. The Microsoft for Windows v2 sensor uses a novel Time-of-Flight (ToF) technology while the previous sensor (Kinect v1) belongs to the category of Structured Light (SL) cameras. Compared with cameras based on SL technology, ToF cameras have a longer range and the images appear to be more accurate without holes in the depth map [62]. The depth map reflects the round-trip time of flight for single laser pulses.

Compared with the previous version, Kinect v2 provides a higher depth map resolution (512×424 vs 320×240), allowing thin objects to be recognized and solving some ambiguity problems. Moreover, Kinect v2 is an inexpensive, unobtrusive and easy to set up sensor that can be used both in home and clinical environments to monitor subjects during physical rehabilitation. The depth features allow the recognition of different subjects and different body parts in the field of view, while the increase in resolution permits the identification of the 3D points of 25 distinct body parts at 30fps. Compared with the previous version, a bio-correction allows each joint to be mapped consistently with an anatomic reference.

C. Preprocessing

In order to filter temporary spikes, a filtering-stage is proposed for the position and orientation of skeleton joints: a 3^{rd} order low-pass Butterworth filter is applied to all the features extracted from the recorded raw data. The cutoff frequency is set at 1Hz according to the residual analysis as described in [63].

D. Exercise Description and Features Extraction

Clinicians selected 5 exercises widely used and clinically recognized for low back pain physiotherapy, providing dynamic dorsal and lumbar stabilization and improving balance in the elderly [64]. The first (Exercise 1) involves the active movement of the upper limbs stretching the trunk muscles, three (Exercises $2-4$) involve active movements of the trunk, one for each of the three space planes, and the last (Exercise 5) involves active movements of the lower limbs. For each exercise, the clinicians specified the primary goals defined as Primary Outcomes (POs), and some constraints, named Control Factors (CFs), in order to map the exercise objectives into kinematic parameters extracted by the 3D joint trajectories. From the absolute quaternion configuration of the Kinect-based motion capture system, it is possible to retrieve the relative quaternions, defined with respect to their parent segment quaternion. This process can be performed following the parent/child multiplications along the quaternion body chain [33]. Subsequently, the conversion of the relative quaternions into Euler angles leads to the derivation of meaningful joint angles (notice that this procedure may lead to the problem of gimbal lock [33]). Goals (POs) and constraints (CFs) became descriptors of the movement, in terms of body segments, distances between anatomical landmarks, and relative angles. Specifically, POs are the target descriptors that change in order to reach the exercise goal (e.g., the maximum range of motion of the upper limbs during their lifting on the frontal plane as for Exercise 3 and the maximum knee flexion on the sagittal plane as in Exercise 5). On the contrary, CFs represent physical constraints which have to be maintained during the exercise (e.g., correct trunk alignment along the sagittal, frontal and transversal plane as in Exercise 2 or stability and complete elbow extension during Exercise 1). In general, correct body alignment during motion is a fundamental requirement for minimizing exercise side effects (pain and muscle contractures) and maximizing the muscle force output during movement. CFs are time series, scalar values that change respect to the time during the exercise execution, while POs are vectors with the same number of elements as the repetitions number and refer to the maximum and minimum of the signal. Both CFs and POs can be considered as vector time series because they are time ordered.

In particular, the reliability and the clinical relevance of these features were explored and confirmed in our recent works [32], [37] by: (i) comparing and validating the clinical features extracted by the Kinect-based motion capture system, and the same features obtained by the stereophotogrammetric system [32], (ii) providing a functional monitoring of these clinical features during exercise execution which disclosed a high level of agreement with respect to clinical judgment [36], [37].

A brief description of each exercise and a related graphical scheme of POs and CFs are reported in Figure 1:

- ***Exercise 1: Lifting of the arms.*** The subject holds a bar with both hands and with arms extended along the body, slightly apart. He/she has to raise the arms above the head, keeping the elbows in extension in order to stretch the trunk muscles. The feet must always be on the ground, slightly apart, with the knees slightly flexed. The subject must avoid anterior or posterior pelvic tilt.
  Extracted Features: angles between right/left arm and upper torso in the sagittal plane ($\alpha_{l/r}$) represent the POs. Elbow extension angles ($\gamma_{l/r}$), knee extension angles ($\phi_{l/r}$), hip angles ($\psi_{l/r}$), torso area (A_t), hands distance (d_h), ankle distance (d_a) are the CFs to be considered.
- ***Exercise 2: Lateral tilt of the trunk with the arms in extension.*** The subject has to raise his/her arms above the head with the elbows completely extended and holding a bar with both hands (starting position). He/she then has to tilt the trunk slowly first to the left and then to

the right, keeping it exactly on the frontal plane. After each tilt, the subject returns to the starting position. The movement must be performed so as to avoid bending the trunk backwards or forwards. The feet must always be on the ground, slightly apart, with the knees slightly flexed. ***Extracted Features:*** right and left angles between the anatomical segment defined by the hip and shoulder and the vertical axis ($\beta_{l/r}$) in the frontal plane (x, y) are defined as POs, while elbow extension ($\gamma_{l/r}$), knee extension angles ($\phi_{l/r}$), hip angles ($\psi_{l/r}$), hand distance (d_h), shoulder distance (d_s), hip distance (d_{hip}) and the vertical distance between the wrists and the shoulders ($h_{l/r}$) and the transverse plane coordinates of the hip ($z_{h_{l/r}}, X_{h_{l/r}}$) normalized to zero mean, are the CFs.

- ***Exercise 3: Trunk rotation.*** The subject holds the arms parallel, at an angle of ninety degrees with respect to the torso (arms aligned with the shoulders) with the elbows completely extended (starting position). He/she then rotates the torso slowly first to the left and afterwards to the right. After rotation to the right, the subject returns to the starting position. During the exercise, the body must be kept well aligned, avoiding bending the trunk backwards or forwards. The feet must always be on the ground, slightly apart, with the knees slightly flexed.
  Extracted Features: PO is the horizontal distance between the elbows (d_x), normalized with respect to the maximum variation. The elbow extension angle ($\gamma_{l/r}$), shoulder extension angles ($\eta_{l/r}$), knee extension angles ($\phi_{l/r}$), hip angles ($\psi_{l/r}$), shoulder distance (d_s), hip distance (d_h) the distance between the wrists and the shoulders ($h_{l/r}$) and the depth coordinates of the hip ($z_{h_{l/r}}$) normalized to zero mean, are the CFs.
- ***Exercise 4: Pelvis rotations on the transverse plane.*** The subject has to stand still with feet slightly apart. Without moving the feet, he/she makes a circular rotation with the pelvis, first in clockwise and then, in the counter clockwise direction.
  Extracted Features: POs are given by the spine base trajectories, normalized to zero mean, in the transverse plane (x, z), to ensure that the subject's position is independent from the sensor. The shoulder distance (d_s), hip distance (d_h), elbow extension ($\gamma_{l/r}$), knee extension angles ($\phi_{l/r}$) and the depth coordinates of the shoulders ($z_{s_{l/r}}$) normalized to zero mean, are the CFs.
- ***Exercise 5: Squatting.*** The subject holds the arms, aligned with the shoulders, at 90° with respect to the trunk with the elbow completely extended (starting position). He/she has to flex the knees up to 60°/70° and then return to the starting position. During the exercise, the body has to be kept well aligned in the sagittal plane so as to avoid bending the trunk backwards or forwards.
  Extracted Features: the right and left knee angles in the sagittal plane ($\theta_{l/r}$) are POs. Hand distance (d_h), shoulder distance (d_s), hip distance (d_{hip}), knee distance (d_k), ankle distance (d_a), torso area (A_t), distance between hand and shoulder ($d_{s_{l/r}}$) and the transverse plane coordinates of the shoulder ($z_{s_{l/r}}, x_{s_{l/r}}$) normalized to zero mean, are the CFs.

EXERCISE ACCURACY ASSESSMENT QUESTIONNAIRE

Taking into account the description and aim of the exercise and observing the whole exercise (all repetitions), please answer the questions choosing one of the following options:

1=Never	2=Rarely	3=Sometimes	4=Often	5=Always

1) Is the primary objective of the exercise achieved (i.e., extension of the upper limbs, trunk rotation with upper limbs elevated to 90°, squatting, etc.) ?
2) Is the exercise repeated in a constant manner?
3) Is the amplitude of the movement complete?
4) Is the posture of the head correct?
5) Is the posture of the right arm correct?
6) Is the posture of the left arm correct?
7) Is the posture of the trunk correct?
8) Is the posture of the pelvis correct?
9) Is the posture of the right leg correct?
10) Is the posture of the left leg correct?

Fig. 2. Clinical assessment questionnaire.

Subjects were asked to repeat each exercise consecutively 5 times and they were positioned at a distance of 3 meters in front of the Kinect sensor, so that distances and angles were calculated in the frontal and sagittal plane, respectively. Each exercise started with the subject in the upright posture with the legs slightly apart. The sequence of the exercises was random.

E. Clinical Assessment

Clinicians, experts in musculoskeletal and neurological disorders (M.C. and M.G.C.), assessed each exercise proposed. They observed videos and compiled the 10-item Likert questionnaire presented in [37], called the Exercise Accuracy Assessment Questionnaire (EAAQ) as reported in Figure 2. This questionnaire was created to quantify the clinicians' judgment as regards the accuracy of the subjects while performing a motor exercise. The first three questions investigated accuracy with respect to the exercises' primary objectives (i.e., extension of the upper limbs, trunk rotation with upper limbs elevated to 90°, squatting, etc.), whereas the last seven items controlled the posture of seven body segments (head/neck, trunk, arms, pelvis and legs), that subjects have to maintain during the exercise. The tool provides three scores: the clinical Total Score (cTS), that is the sum of the ten identified scores; the clinical Primary Outcome (cPO) score, that is the sum of the scores of the first three questions; and, finally, the clinical Control Factors (cCF) as the sum of the last seven items (about postural performance).

F. Data Description

The dataset description is shown in Figure 3. The enrolled population is presented and split into the two previously defined macro-groups: the *Control Group* (CG) and the group of people with *Pain and Posture disorders* (GPP). The CG is subdivided into two subgroups with (*CG-E*) or without (*CG-NE*) expertise in physiotherapy exercises, while the *GPP* is divided into 3 sub-groups according to the diagnosis: *Stroke* (*GPP-S*), *Parkinson's disease* (*GPP-P*) and *Low Back Pain* (*GPP-B*). In each group, the subjects have their own folder with all the exercises performed. For each of the 5 exercises, the authors provided 3 sub-folders related to the *Raw* data, the *Script* and the *Label* as follows:

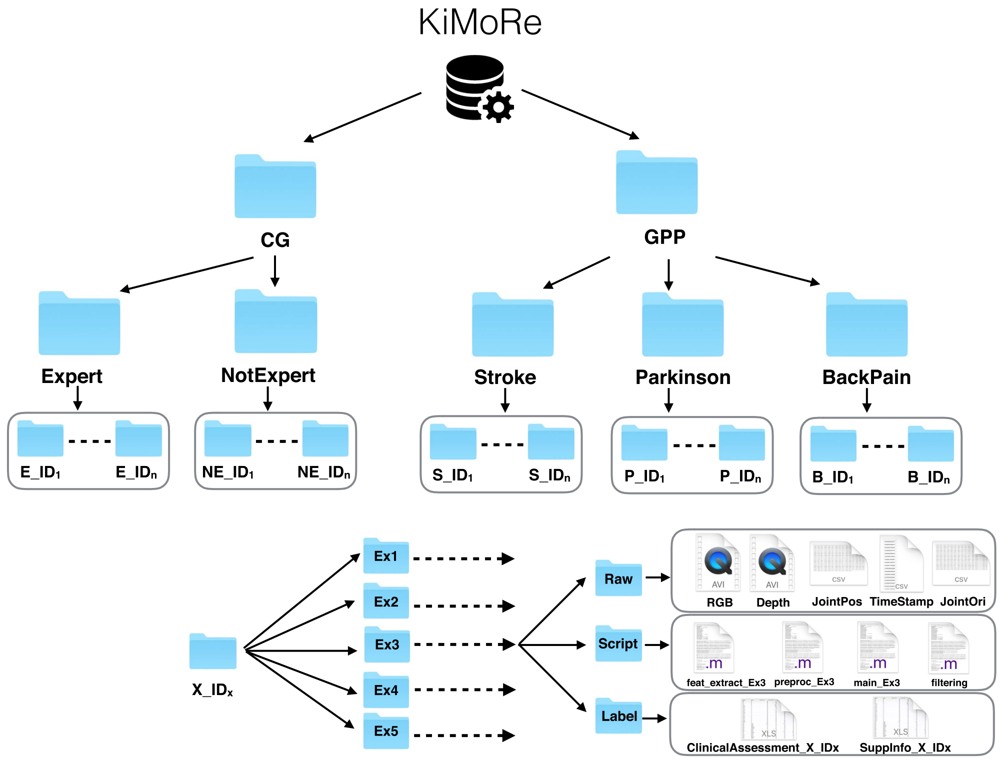


Fig. 3. Schematic description of data presentation in the proposed dataset structure.

- the *Raw* folder includes raw data acquired directly from the Kinect v2 sensor that are related to the RGB video, depth video, the joint positions and orientations, and the time stamp with the acquisition times. Respectively, the files available in this folder are *(depthDDMMYY_XXXXXX)*, *(JointPositionDDMMYY_XXXXXX)*, *(JointOrientationDDMMYY_XXXXXX)* and *(TimeStampDDMMYY_XXXXXX)*, where *DDMMYY* refers to the acquisition date and *XXXXXX* are associated numbers for each recording;
- the *Script* folder includes the code related to the implemented functions, in particular the features extraction step and the pre-processing of data, both called back in the main function. The code for data filtering is also available in this folder. The related files are *feat_extract_ExX*, *preproc_ExX*, *main_ExX*, and *filtering*;
- the *Label* folder includes two files: *ClinicalAssessment_X_IDx*, related to the clinical scores assigned by clinicians, including both total and local scores, and *SuppInfo_X_IDx* that provides information about sex, age, diseases and other supplementary information that might affect the subject.

RGB data will be made available on explicit request to the corresponding author and after signing an End User License Agreement (EULA document).

III. Results

The clinical impact and the reliability of the KIMORE dataset were measured by (i) validating the accuracy of the recorded data with respect to a gold-standard system, (ii) testing the clinical validity of the questionnaire and (iii) demonstrating how this dataset has potential to be used to build a template/rule-based model for evaluating the patient's performance during rehabilitation. In particular, we provide evidence of how the KIMORE dataset can be used to train, validate and test human motion assessment approaches.

A. Validation of the Dataset

The validation of the dataset was performed according to our published study [32] considering two of four recorded exercises involving the upper body (i.e., Exercises 1 and 2) and one involving the lower body (i.e., Exercise 5). The accuracy of the Kinect motion capture was measured in terms of spatial and temporal accuracy with respect to the gold standard, represented by a stereophotogrammetric system (ELITE, BTSEngineering, Milano), characterized by 6 infrared cameras. In particular, the validation study is provided by analyzing (i) raw data provided directly by Kinect skeleton tracking (i.e., joint displacement), (ii) the POs and (iii) the temporal difference between the POs computed by Kinect and the gold-standard system. The results of the experimental validation [32] provide evidence of the consistency of features extracted by Kinect motion capture data, with respect to the gold standard motion capture. The analysis of raw data was performed considering the displacement between the elbow joints for Exercise 1 and Exercise 2 and the displacement between the ankle joints for Exercise 5. This analysis demonstrates how Kinect v2 follows the trend of the gold standard. However, there is a static offset for the same joint (i.e., 8cm and 3cm for elbow and ankle joints) which could be easily removed in order to track the performance of the patient over time.

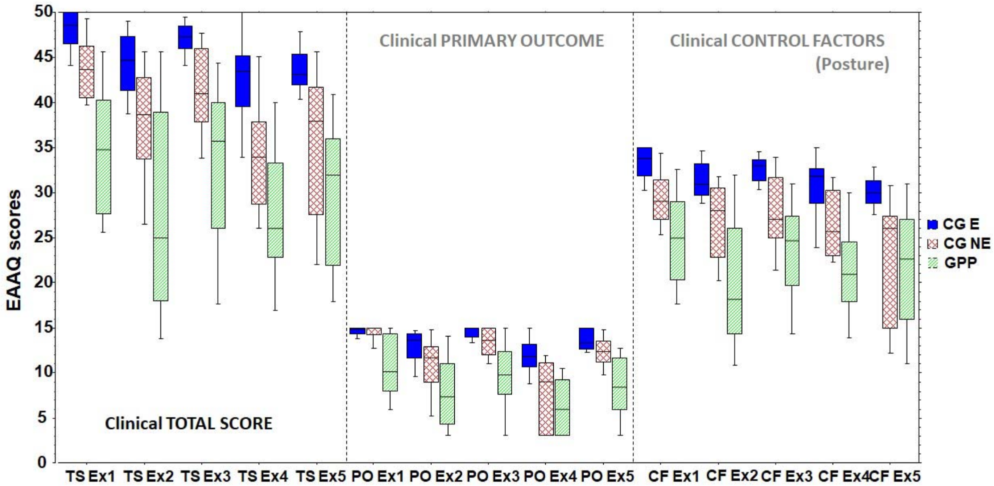


Fig. 4. Legend: CG-E = Control Group- Experts; CG-NE= Control Group-Non-Experts; GPP X Group of people with Pain and Postural disorders; TS Ex 1-5 = clinical Total Score Exercises 1-5: PO Ex 1-5 = clinical Primary Outcome Exercises 1-5; CF Ex1-5 = clinical Control Factors Exercises 1-5.

The Microsoft Kinect sensor seems to be more reliable for tracking POs in the motor task involving the upper limbs (i.e., Exercise 1: maximum relative error = 12.1%) with respect to the task involving the lower body (i.e., Exercise 3: maximum relative error = 26.3%). However, during maximum knee flexion in Exercise 5 there is a systematic bias between the two measurements: some main Kinect joints used for computing the POs of Exercise 5 can be occluded. On the other hand for the exercises involving the upper body (e.g., Exercise 1) the larger measurement volume can increase error variability (up to 9.5°). For Exercise 2 there is a vertical symmetry across the frontal plane: the error is comparable during left and right oscillation (i.e., relative error = 12% and 12.7% respectively).

The temporal accuracy results confirm how Kinect v2 could accurately measure the timing characteristic of exercises (maximum absolute latency 0.03 ± 0.113 sec).

The performed comparison is shown in greater detail in Appendix A (see Table IV).

B. Clinical Questionnaire Validation

Questionnaire validity was tested on normal and pathological subjects where it proved to be able to distinguish between healthy and disabled people [36], while the inter-rater reliability was checked, comparing the judgments of three clinicians (one physician and two physiotherapists) applying Cohen's Kappa test which reached a K-value > 0.8 [37].

Figure 4 shows the box plot of the EAAQ scores as well as the mean, standard deviation and standard errors of the three groups (CG-E, CG-NE and GPP) for the five exercises.

The differences between groups were analyzed applying the Kruskall Wallis test and the results are detailed in Table II: the clinical total score of the GPP was significantly lower than the CG, where the highest scores were achieved by the experts. Since it is also important to prove that the questionnaire is able to distinguish patients from people without any expertise in physiotherapy exercises, a direct comparison between GPP and CG-NE was carried out applying the Mann-Whitney U test. The three clinical scores, i.e. cTS, cPO and cCF, differed significantly between the two groups for each of the five exercises: in fact, the comparison reached a Z score > 4.00 and a p value < .0001 in all cases except for cCF for Exercise 5 ($Z = -3.1$ and $p = .002$), cPO for Exercise 4 ($Z = -3.6$ and $p = .0008$) and cPO for Exercise 1 ($Z = -3.9$ and $p = .0001$), where, nonetheless, significant differences between groups where highlighted. The figure also shows that cPO and cCF of GPP subjects differ from those of CG. These results highlighted, therefore, that a template-based approach should be based on the movement performed by experts in movement therapy.

TABLE II
COMPARISON BETWEEN GROUPS (CG-E VS CG-NE VS GPP) OBTAINED WITH THE KRUSKALL WALLIS TEST. THE H AND THE P VALUE ARE REPORTED FOR EACH EXERCISE AND FOR cTS, cPO AND cCF

	cTS (H; p value)	cPO (H; p value)	cCF (H; p value)
Ex. 1	34.8; < .0001	20.4; < .0001	30.1; < .0001
Ex. 2	24.9; < .0001	20.1; < .0001	25.0; < .0001
Ex. 3	37.3; < .0001	27.4; < .0001	29.7; < .0001
Ex. 4	31.8; < .0001	21.4; < .0001	32.2; < .0001
Ex. 5	27.3; < .0001	35.7; < .0001	17.4; = .0002

C. The Usefulness of the Dataset in a Rehabilitation Assessment Scenario

In the present paper, we tested, on a large population (n = 78), the correlation of the EAAQ total score with respect to the assessment performed through the instrumental rule-based methodology proposed in [37] and the template-based methodology proposed in [36]. A Spearman rank correlation test [65] was applied for this scope and the analysis results are displayed in Table II and Table III.

In addition, in Appendix B we have provided a full example of how the dataset can be used to evaluate the performance objectively using a rule- or template-based approach.

TABLE III

CORRELATION ANALYSIS BETWEEN THE EAAQ AND THE SCORES OBTAINED BY THE RULE AND TEMPLATE-BASED ALGORITHMS APPLYING THE SPEARMAN CORRELATION: Z, p AND ρ VALUES ARE REPORTED IN THE TABLE. THE SIGNIFICANCE, WAS SET AT .05 FOR ALL CORRELATION COMPARISON. HENCE, N.S. INDICATES THAT THE CORRELATION RESULT IS NOT STATISTICALLY SIGNIFICANT

	rule-based			template-based
	TS	PO	CF	TS
Ex. 1	3.3; .001; .4	3.3; .002; .4	n.s.	2.6; .01; 0.4
Ex. 2	2.7; .006; .4	2.6; .008; .4	2.6; .008; .4	2.6; < .0001; 0.7
Ex. 3	2.8; .005; .4	3.6; .0004; .5	2.5; .01; .3	2.9; .007; 0.5
Ex. 4	4.2; < .0001; .6	2.7; .007; .4	4.2; .002; .4	2.4; .02; .4
Ex. 5	2.2; .03; .3	3.3; .0009; .5	n.s.	n.s.

IV. DISCUSSION

The paper presents the KIMORE dataset including 5 exercises, which are widely used for posture and back pain rehabilitation, collected using a RGB-D based skeleton tracking system, in a clinical outpatient scenario. The KIMORE includes 78 subjects with (34) and without (44) neurological or musculoskeletal disorders, affected by postural disturbances or back pain. KIMORE is comprised of RGB, depth and 25 joint positions and orientations. In addition, the total and local scores, provided by the medical staff, according to the clinical assessment questionnaire published in [37], are available for each exercise.

The main contributions of the introduced KIMORE dataset compared to the related literature are:

- the high number and heterogeneity of enrolled subjects with respect to the literature;
- the organization of collected data in different groups on the basis of diagnosis or expertise;
- the collaborative approach between engineers and clinicians in designing the experimental procedure;
- the identification by clinicians of two main groups of features to monitor, defined as primary outcomes and postural constraints;
- the accurate description of the main motor task features together with a specific algorithm for their extraction, available with the Matlab code.
- the annotation of the dataset carried out by two expert clinicians according to a questionnaire validated in [37], related to the achievement of the primary outcomes and kinematic constraints of each exercise, is included.
- KIMORE reports core exercises useful in widespread pathological conditions (i.e., back pain and postural disturbances) [34], [35] providing a detailed dataset for rehabilitation subjects of all ages and socioeconomic status who seek health care [66]. Although the present study was run at a hospital facility, in order to respond to validation needs, the architecture was built to be easily delivered at home.

All these aspects support the detailed description of the proposed work by providing a method which is useful for building telerehabilitation systems. The RGB-based telerehabilitation system has been demonstrated to support exercise therapy, by showing similar or better effectiveness compared to a conventional therapy [61]. Nevertheless, more research is advocated to provide insight into motion analysis for musculoskeletal rehabilitation, because of the low methodological quality of the reviewed studies.

The high number of enrolled subjects (total 78 subdivided into 44 healthy and 34 with posture and pain disorders) helps the generalization of the results. Data are organized systematically: those related to the 34 subjects constituting the *GPP* are organized in the dataset with respect to the pathology (i.e., low back pain due to spondylosis, hemiparesis due to cerebral stroke, Parkinson's disease). The 44 healthy subjects that make up the *CG* include both experts in physiotherapy (i.e., physiotherapists and physiotherapy students) and people without specific skills in the rehabilitation context.

As the authors highlighted in Section I-A and Table I, the proposed dataset aims to fill a gap with respect to the existing literature, promoting the integration and applicability of rule/template-based models for assessing the performance of rehabilitation programs.

A. The Impact of the KIMORE Dataset for Rehabilitation Assessment

Human motion assessment based on Artificial Intelligence (AI) can be divided into *rule* and *template* according to the scientific literature [17]. Since both these methodologies are powered by data, the main requirement for reliable design and the consequent application of these algorithms is to have a high-quality labeled training dataset. In this context the KIMORE dataset provides, but is not limited, to the following opportunities. It can be used to train a template-based approach (i.e., supervised Machine Learning (ML) model). For example, in [37] the Hidden Semi Markov Model was trained using only the features of the best subjects (cTS) of the control group, considered as an exemplar of the motion sequence. The instrumental score was computed measuring the likelihood of the observation with respect to the trained model. In this scenario the KIMORE dataset opens up a whole range of possibilities which can be used to provide suitable data for training the ML model. From the ML perspective, the trained model may be designed so as to evaluate the subjects' performance for different exercises by generalizing across unseen subjects.

The rule-based approach can exploit the proposed dataset by properly setting the rule parameters (i.e., objective and tolerance value [37] or parameter of the membership function for the fuzzy inference [67]) for each different exercise. This process may be considered as the main core of the collaborative design procedure described in [67].

KIMORE is highly dependent on the collaboration of medical staff, in order to compile the "ground truth label" regarding the accuracy of exercise performance, which is the main gap with respect to the available RGB-D-based action recognition datasets [38]. The medical staff contribution concerns the definition of relevant clinical motion features and a questionnaire designed/validated for exercise performance assessment. The features are extrapolated by the exercise description in terms of the movement and posture that subjects have to adopt during the motor task performance, while clinicians based their judgment on the available literature [34], [64]. The exercise description also provides the reference to fill in a questionnaire (i.e., EAAQ). The EAAQ is a task-independent assessment tool, in which judgment is based on the scope and kinematic characteristics of the motor exercise. To the best of the authors' knowledge, there are no validated clinical tools in literature, to rate individual performance of a therapeutic exercise, although physiotherapists usually supervise patients and give them constant feedback on how to optimize motor performance [68]. The questionnaire may be useful to monitor the exercise quantitatively, to track the rehabilitation functional outcomes and to test and compare different machine learning and rule-based methodologies. Monitoring how an exercise is performed promotes the outcome and avoids possible side effects due to incorrect postures or incomplete movements [61]. Moreover, giving patients feedback during or after the exercise improves their motor learning, thereby enhancing retention of new motor skills [6]. The clinical questionnaire has proved to be able to distinguish between the subgroups of subjects with appropriate reliability [36], [37]. Furthermore, KIMORE presents data of complex gestures while monitoring the whole body posture to control factors that may influence exercise outcomes. This novelty tries to reflect clinical scenarios, so as to correctly validate the method and to increase its generalization. However, the available data proposed are not limited to clinical features and skeleton tracking trajectories. The authors provided the main streams of RGB and depth so that the dataset can be easily generalized for different computer vision tasks without being strictly related to the defined clinical features and the Kinect skeleton tracking algorithms. This aspect, together with the high level of generalization of the clinical features and joint orientation encourages the application of this dataset for modeling, comparing and validating different rule/template-based approaches in the physical rehabilitation scenario. The feature extraction and the questionnaire may be generalized to any exercise. The clinicians provided the cTS, the cPO score and the cCF score in order to quantify the overall performance of the subjects and to evaluate the achievement of the primary objectives and the postural constraints, respectively. In this context, the reliability of template/rule-based (instrumental) approaches can be measured according to the correlation between the clinical scores (i.e., cTS, cPO and cCF) and scores provided by the instrumental approach for different sets of exercises. As evidence of this concept, we performed a non-linear correlation analysis between one template-based algorithm proposed in [36], and one rule-based algorithm introduced in [37] with respect to the ground-truth scores (see Section III-C). In particular, the correlation, with respect to the clinicians' judgments, highlights a moderately significant reliability of TS and PO scores computed by the rule-based approach, while the CF scores are less accurate for assessing postural features. However, the moderate correlation value (around 0.4) shows how the solution of the rehabilitation assessment task is not a trivial problem. In this scenario, the rule-based and template-based approaches can prevail over each other in order to better evaluate human movement. In fact the template-based approach is better than the rule-based approach for Exercises 2 and 3, while the rule-based approach seems to be more accurate for Exercises 4 and 5. The low performance of Exercise 5 for both methodologies can be justified by some limitations of the RGB-D sensor, which is also confirmed by the validation analysis (see Section III-A). Dynamic movements may include different postures adverse to the vision sensor characterized by joint occlusion [32]. In this context, the skeletal tracking algorithm seems to ensure a better accuracy in the motor task involving the upper limbs (e.g., Exercises 1 and 2) with respect to the task involving the lower body (Exercise 5). However, neither approach is constantly able to capture the same components as detected by a clinician who is monitoring the movement. Moreover, the clinical assessment can also be affected by bias and errors due to the discrete EAAQ scale (e.g., rarely vs sometimes vs often) and to the visual inspection of a 3D human movement performed through 2D video images. These elements and the sample heterogeneity can affect the variability of the clinical measurements (see Figure 4).

V. Conclusions

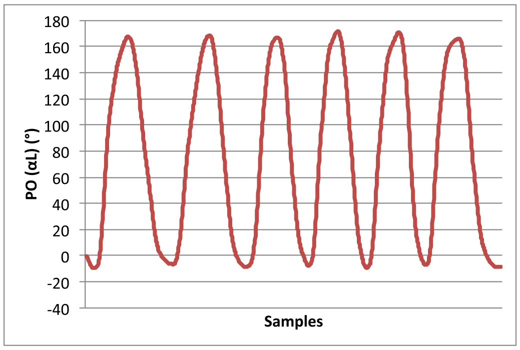
KIMORE may be useful for building a remote rehabilitation system for low back pain therapy and posture disability, to meet both health care and patients' needs for the continuity and sustainability of health care services for chronic disabilities. Future works are warranted to study a greater sample of exercises and subjects so that, even if bigger than other datasets, KIMORE may be augmented in order to model the high variability of human movement. In this study, the selected exercises were chosen on the basis of their reproducibility in front of the Kinect sensor. Other types of exercises that present a partial or total occlusion of the tracked joints involved in the movement should be explored in order to measure the accuracy of the sensor.

Appendix A
Validation of the Dataset

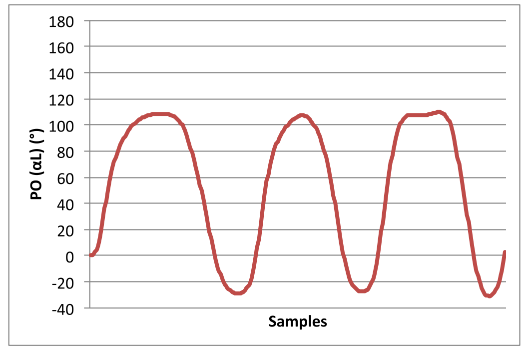
Table IV shows the results of the KIMORE validation with respect to the gold standard stereophotogrammetric system. For more detail, please refer to the published paper [32].

Appendix B
The Usefulness of the Dataset in a Rehabilitation Assessment Scenario

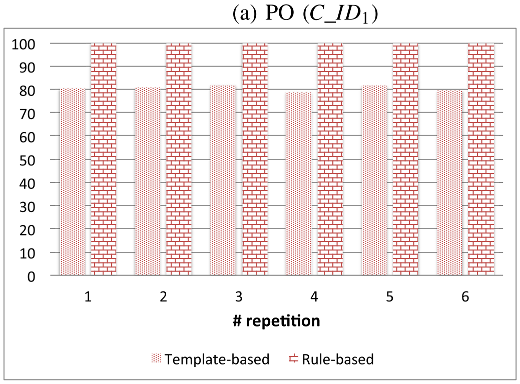
An example of how the dataset can be used for rehabilitation assessment is reported below. We have applied our rule- and template-based approaches published respectively in [37] and [36], exploiting the potential of the presented KIMORE dataset, and the physical exercise assessment was performed according to these methodologies. As regards the rule-based approach, we set the rule parameters (i.e., objective and tolerance value) according to the value of the CG subjects. On the contrary, the Hidden Semi-Markov Model (HSMM) was applied to evaluate body motion during a rehabilitation training program. The training of the HSMM was carried out using only the features of a subset of CG who achieved the highest cTS. The chosen features are collected and available within the KIMORE dataset (i.e., PO and CF described in Section II-C). Both the approaches considered are able to



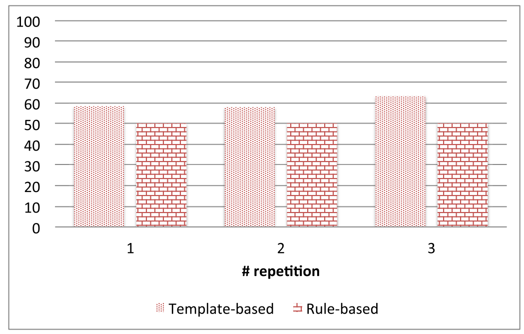
(a) PO (C_ID_1)



(b) PO (P_ID_{26})

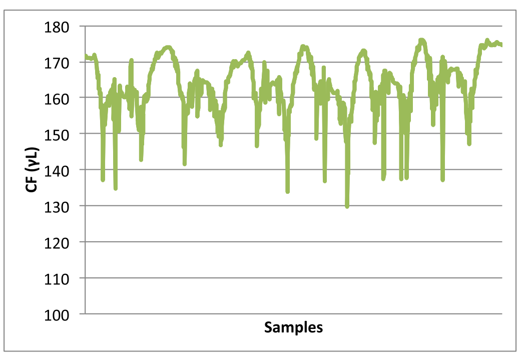


(c) PO scores (C_ID_1)

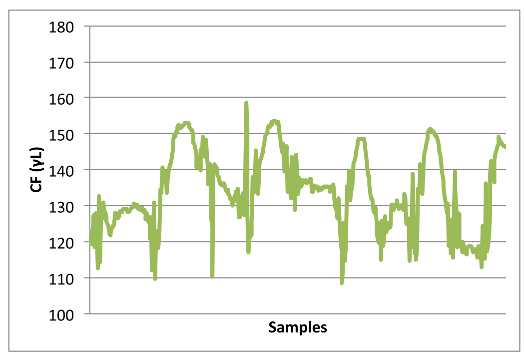


(d) PO scores (P_ID_{26})

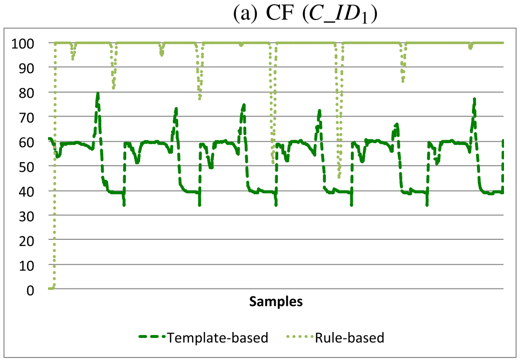
Fig. 5. PO (α_L) of C_ID_1 (a) and P_ID_{26} (b) for Exercise 1. The relative PO scores (c) and (d) were computed for each repetition according to the rule- [37] and template-based approach [36].



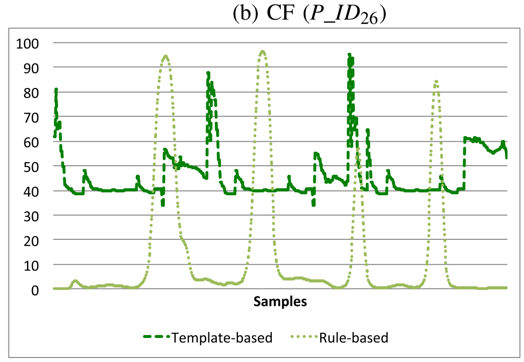
(a) CF (C_ID_1)



(b) CF (P_ID_{26})



(c) CF scores (C_ID_1)



(d) CF scores (P_ID_{26})

Fig. 6. CF (γ_L) of C_ID_1 (a) and P_ID_{26} (b) for Exercise 1. The relative CF scores (c) and (d) were computed for each time stamp according to the rule- [37] and template-based approach [36].

provide disaggregated and total scores. Since the PO aims to evaluate the achievement of the maximum and minimum target angle/position, the PO related scores are generally computed for each repetition of the considered exercise. Figure 5 shows the PO scores (relative to α_L) computed for each repetition of Exercise 1 for C_ID_1 and P_ID_{26}.

On the contrary the CF describes a constraint achievement over time (e.g., subjects have to keep the elbow extended to

TABLE IV
COMPARISON BETWEEN KINECT AND THE GOLD STANDARD STEREOPHOTOGRAMMETRIC SYSTEM

	Raw data comparison: joint displacements		
	Joints	Offset (cm)	Root Mean Square Error (cm)
Ex. 1	elbows	8.3	2.7
Ex. 2	elbows	7.8	4.7
Ex. 5	ankles	3	2.4
	PO comparison		
	PO	Absolute error (°)	Relative error (%)
Ex. 1	max(α_r)	18.0 ± 9.5	12.1
	min(α_r)	11.4 ± 6.4	7.5
	max(α_l)	13.1 ± 9.4	8.7
	min(α_l)	8.1 ± 5.3	5.3
Ex. 5	max(θ_r)	9.5 ± 5.9	9.7
	min(θ_r)	24 ± 10.4	24.3
	max(θ_l)	8.2 ± 2.7	8.4
	min(θ_l)	26 ± 8.1	26.3
	Temporal synchronization		
	PO	Absolute error (# frames)	
Ex. 1	α_r	1.4 ± 3.4	
	α_l	1.0 ± 3.4	
Ex. 5	θ_r	0.5 ± 1.1	
	θ_l	0.5 ± 2.1	

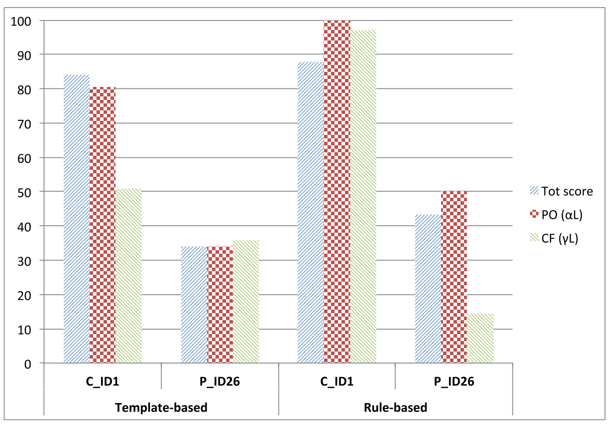


Fig. 7. Overall PO (α_L), CF (γ_L) and total scores for C_ID_1 and P_ID_{26} during Exercise 1.

about 180° over time for Exercise 1). Thus the CF scores were extracted for each timestamp. Figure 6 shows the CF scores related to α_L during Exercise 1 for C_ID_1 and P_ID_{26}.

The overall PO (α_L), CF (γ_L) and total scores are reported in Figure 7 for Exercise 1.

The total score, computed according to the rule- and template-based approach, is respectively 84 and 87 for C_ID_1 and 34 and 43 for P_ID_{26}. These scores are in line with the clinical questionnaire which indicates a cTS of 98 for C_ID_1 and a cTS of 34 for P_ID_{26}. In addition to the total score, the authors provide disaggregated scores for the PO and CF features involved that allow the clinician to localize the error in the exercise movement execution. The reliability of the rule- and template- based approach is not limited to measuring the correlation between the cTS and the computed total score. The computed PO and CF scores can be compared with respect to the cPO and cCF (see Section III-C for more details on this performed comparison).

ACKNOWLEDGMENTS

The authors would like to thank the physiotherapists Paola Casoli and Martina Grugnetti for their help during the clinical assessment procedure.

REFERENCES

[1] G. Barry, B. Galna, and L. Rochester, "The role of exergaming in Parkinson's disease rehabilitation: A systematic review of the evidence," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, p. 33, 2014.

[2] D. Hailey, R. Roine, A. Ohinmaa, and L. Dennett, "Evidence of benefit from telerehabilitation in routine care: A systematic review," *J. Telemed. Telecare*, vol. 17, no. 6, pp. 281–287, 2011.

[3] D. Kairy, P. Lehoux, C. Vincent, and M. Visintin, "A systematic review of clinical outcomes, clinical process, healthcare utilization and costs associated with telerehabilitation," *Disab. Rehabil.*, vol. 31, no. 6, pp. 427–447, 2009.

[4] M. Rogante, M. Grigioni, D. Cordella, and C. Giacomozzi, "Ten years of telerehabilitation: A literature overview of technologies and clinical applications," *NeuroRehabilitation*, vol. 27, no. 4, pp. 287–304, 2010.

[5] K. Steel, D. Cox, and H. Garry, "Therapeutic videoconferencing interventions for the treatment of long-term conditions," *J. Telemed. Telecare*, vol. 17, no. 3, pp. 109–117, 2011.

[6] M. R. Albuquerque, G. M. Lage, H. Ugrinowitsch, U. C. Corrêa, and R. N. Benda, "Effects of knowledge of results frequency on the learning of generalized motor programs and parameters under conditions of constant practice," *Perceptual Motor Skills*, vol. 119, no. 1, pp. 69–81, 2014.

[7] L. O. Costa *et al.*, "Motor control exercise for chronic low back pain: A randomized placebo-controlled trial," *Phys. Therapy*, vol. 89, no. 12, pp. 1275–1286, 2009.

[8] M. van Diest, C. J. C. Lamoth, J. Stegenga, G. J. Verkerke, and K. Postema, "Exergaming for balance training of elderly: State of the art and future developments," *J. Neuroeng. Rehabil.*, vol. 10, no. 1, p. 101, 2013.

[9] M. van Diest *et al.*, "Suitability of Kinect for measuring whole body movement patterns during exergaming," *J. Biomechanics*, vol. 47, no. 12, pp. 2925–2932, 2014.

[10] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea, "Machine recognition of human activities: A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 11, pp. 1473–1488, Nov. 2008.

[11] R. Poppe, "A survey on vision-based human action recognition," *Image Vis. Comput.*, vol. 28, no. 6, pp. 976–990, 2010.

[12] L. Pogrzeba, T. Neumann, M. Wacker, and B. Jung, "Analysis and quantification of repetitive motion in long-term rehabilitation," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 3, pp. 1075–1085, May 2018.

[13] B. Dehbandi *et al.*, "Using data from the Microsoft Kinect 2 to quantify upper limb behavior: A feasibility study," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 5, pp. 1386–1392, Sep. 2017.

[14] I. Ar and Y. S. Akgul, "A computerized recognition system for the home-based physiotherapy exercises using an RGBD camera," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 6, pp. 1160–1171, Nov. 2014.

[15] V. Bonnet and G. Venture, "Fast determination of the planar body segment inertial parameters using affordable sensors," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 4, pp. 628–635, Jul. 2015.

[16] R. J. Adams *et al.*, "Virtual activities of daily living for recovery of upper extremity motor function," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 1, pp. 252–260, Jan. 2018.

[17] W. Zhao, R. Lun, D. D. Espy, and M. A. Reinthal, "Rule based realtime motion assessment for rehabilitation exercises," in *Proc. IEEE Symp. Comput. Intell. Healthcare e-Health (CICARE)*, Dec. 2014, pp. 133–140.

[18] M. Capecci *et al.*, "A tool for home-based rehabilitation allowing for clinical evaluation in a visual markerless scenario," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 8034–8037.

[19] A. K. Mishra, M. Skubic, and C. Abbott, "Development and preliminary validation of an interactive remote physical therapy system," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 190–193.

[20] S. Lee, Y.-S. Lee, and J. Kim, "Automated evaluation of upper-limb motor function impairment using Fugl-Meyer assessment," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 1, pp. 125–134, Jan. 2018.

[21] M.-C. Hu, C.-W. Chen, W.-H. Cheng, C.-H. Chang, J.-H. Lai and J.-L. Wu, "Real-time human movement retrieval and assessment with Kinect sensor," *IEEE Trans. Cybern.*, vol. 45, no. 4, pp. 742–753, Apr. 2015.

[22] C.-J. Su, C.-Y. Chiang, and J.-Y. Huang, "Kinect-enabled home-based rehabilitation system using dynamic time warping and fuzzy logic," *Appl. Soft Comput.*, vol. 22, pp. 652–666, Sep. 2014.

[23] W. Ding, K. Liu, X. Fu, and F. Cheng, "Profile HMMs for skeleton-based human action recognition," *Signal Process., Image Commun.*, vol. 42, pp. 109–119, Mar. 2016.

[24] J. F.-S. Lin and D. Kulic, "Online segmentation of human motion for automated rehabilitation exercise analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 1, pp. 168–180, Jan. 2014.

[25] L. Ciabattoni, F. Ferracuti, S. Iarlori, S. Longhi, and L. Romeo, "A novel computer vision based e-rehabilitation system: From gaming to therapy support," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2016, pp. 43–44.

[26] M. Nieto-Hidalgo, F. J. Ferrández-Pastor, R. J. Valdivieso-Sarabia, J. Mora-Pascual, and J. M. García-Chamizo, "A vision based proposal for classification of normal and abnormal gait using RGB camera," *J. Biomed. Inform.*, vol. 63, pp. 82–89, Oct. 2016.

[27] D. Leightley, J. S. McPhee, and M. H. Yap, "Automated analysis and quantification of human mobility using a depth sensor," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 4, pp. 939–948, Jul. 2017.

[28] I. Pernek, G. Kurillo, G. Stiglic, and R. Bajcsy, "Recognizing the intensity of strength training exercises with wearable sensors," *J. Biomed. Inform.*, vol. 58, pp. 145–155, Dec. 2015.

[29] Z. Zhang, Q. Fang, and X. Gu, "Objective assessment of upper-limb mobility for poststroke rehabilitation," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 859–868, Apr. 2016.

[30] D. Antón, A. Goñi, A. Illarramendi, J. J. Torres-Unda, and J. Seco, "KiReS: A Kinect-based telerehabilitation system," in *Proc. IEEE 15th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Oct. 2013, pp. 444–448.

[31] J. Shotton *et al.*, "Efficient human pose estimation from single depth images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2821–2840, Dec. 2013.

[32] M. Capecci *et al.*, "Accuracy evaluation of the Kinect v2 sensor during dynamic movements in a rehabilitation scenario," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 5409–5412.

[33] A. Napoli, S. Glass, C. Ward, C. Tucker, and I. Obeid, "Performance analysis of a generalized motion capture system using microsoft kinect 2.0," *Biomed. Signal Process. Control*, vol. 38, pp. 265–280, Sep. 2017.

[34] C. Kisner and L. Colby, *Therapeutic Exercise: Foundations and Techniques*. Philadelphia, PA, USA: F.A. Davis, 2012.

[35] M. Hutson and A. Ward, *Oxford Textbook of Musculoskeletal Medicine*. Oxford, U.K.: Oxford Univ. Press, 2015.

[36] M. Capecci *et al.*, "A hidden semi-Markov model based approach for rehabilitation exercise assessment," *J. Biomed. Inform.*, vol. 78, pp. 1–11, Feb. 2018.

[37] M. Capecci *et al.*, "An instrumental approach for monitoring physical exercises in a visual markerless scenario: A proof of concept," *J. Biomechanics*, vol. 69, pp. 70–80, Mar. 2018.

[38] J. Zhang, W. Li, P. O. Ogunbona, P. Wang, and C. Tang, "RGB-D-based action recognition datasets: A survey," *Pattern Recognit.*, vol. 60, pp. 86–105, Dec. 2016.

[39] J. Shin *et al.*, "Open access dataset for EEG+ NIRS single-trial classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1735–1745, Oct. 2017.

[40] Y. Wang, X. Chen, X. Gao, and S. Gao, "A benchmark dataset for SSVEP-based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1746–1752, Oct. 2017.

[41] W. L. Lim, O. Sourina, and L. P. Wang, "STEW: Simultaneous task EEG workload data set," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2106–2114, Nov. 2018.

[42] G. Ciocca, P. Napoletano, and R. Schettini, "Food recognition: A new dataset, experiments, and results," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 3, pp. 588–598, May 2017.

[43] W. Li, Y. Zhao, X. Chen, Y. Xiao, and Y. Qin, "Detecting Alzheimer's disease on small dataset: A knowledge transfer perspective," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 3, pp. 1234–1242, May 2018.

[44] Z. Cai, J. Han, L. Liu, and L. Shao, "RGB-D datasets using microsoft Kinect or similar sensors: A survey," *Multimedia Tools Appl.*, vol. 76, no. 3, pp. 4313–4355, Feb. 2017.

[45] C. Chen, R. Jafari, and N. Kehtarnavaz, "UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 168–172.

[46] S. Khalighi, T. Sousa, J. M. Santos, and U. Nunes, "ISRUC-Sleep: A comprehensive public dataset for sleep researchers," *Comput. Methods Programs Biomed.*, vol. 124, pp. 180–192, Feb. 2016.

[47] H. P. Da Silva, A. Lourenço, A. Fred, N. Raposo, and M. Aires-de-Sousa, "Check your biosignals here: A new dataset for off-the-person ECG biometrics," *Comput. Methods Programs Biomed.*, vol. 113, no. 2, pp. 503–514, 2014.

[48] M. Edwards, J. Deng, and X. Xie, "From pose to activity: Surveying datasets and introducing CONVERSE," *Comput. Vis. Image Understand.*, vol. 144, pp. 73–105, Mar. 2016.

[49] H. Wu, X. Ma, Z. Zhang, H. Wang, and Y. Li, "Collecting public RGB-D datasets for human daily activity recognition," *Int. J. Adv. Robot. Syst.*, vol. 14, no. 4, 2017, Art. no. 1729881417709079.

[50] M. S. H. Aung *et al.*, "The automatic detection of chronic pain-related expression: Requirements, challenges and the multimodal EmoPain dataset," *IEEE Trans. Affect. Comput.*, vol. 7, no. 4, pp. 435–451, Oct./Dec. 2016.

[51] D. J. Rude, S. Adams, and P. A. Beling, "A benchmark dataset for depth sensor based activity recognition in a manufacturing process," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 668–674, 2015.

[52] W. Zhang, Z. Liu, L. Zhou, H. Leung, and A. B. Chan, "Martial arts, dancing and sports dataset: A challenging stereo and multi-view dataset for 3D human pose estimation," *Image Vis. Comput.*, vol. 61, pp. 22–39, May 2017.

[53] L. Chen, H. Wei, and J. Ferryman, "ReadingAct RGB-D action dataset and human action recognition from local features," *Pattern Recognit. Lett.*, vol. 50, pp. 159–169, Dec. 2014.

[54] B. Ni, G. Wang, and P. Moulin, "RGBD-HuDaAct: A color-depth video database for human daily activity recognition," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, vol. 47, Nov. 2011, pp. 1147–1153.

[55] A. Vakanski, H.-P. Jun, D. Paul, and R. Baker, "A data set of human body movements for physical rehabilitation exercises," *Data*, vol. 3, no. 1, p. 2, 2018.

[56] F. Negin, F. Özdemir, C. B. Akgül, K. A. Yüksel, and A. Erçil, "A decision forest based feature selection framework for action recognition from RGB-depth cameras," in *Image Analysis and Recognition*, M. Kamel and A. Campilho, Eds. Berlin, Germany: Springer, 2013, pp. 648–657.

[57] D. Leightley, M. H. Yap, J. Coulson, Y. Barnouin, and J. S. McPhee, "Benchmarking human motion analysis using Kinect one: An open source dataset," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA)*, 2015, pp. 1–7.

[58] A. Bernardino *et al.*, "A dataset for the automatic assessment of functional senior fitness tests using kinect and physiological sensors," in *Proc. 1st Int. Conf. Technol. Innov. Sports, Health Wellbeing (TISHW)*, Dec. 2016, pp. 1–6.

[59] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in *Proc. CVPR*, 2017, pp. 7291–7299.

[60] M. van Middelkoop *et al.*, "Exercise therapy for chronic nonspecific low-back pain," *Best Pract. Res. Clin. Rheumatol.*, vol. 24, no. 2, pp. 193–204, 2010.

[61] J. Verbrugghe *et al.*, "Motion detection supported exercise therapy in musculoskeletal disorders: A systematic review," *Eur. J. Phys. Rehabil. Med.*, vol. 54, no. 4, pp. 591–604, 2018.

[62] L. Chen, H. Wei, and J. Ferryman, "A survey of human motion analysis using depth imagery," *Pattern Recognit. Lett.*, vol. 34, no. 15, pp. 1995–2006, Nov. 2013.

[63] D. Winter, *Biomechanics and Motor Control of Human Movement*. Hoboken, NJ, USA: Wiley, 2009. [Online]. Available: https://books.google.it/books?id=_bFHL08IWfwC

[64] M. J. Vezina and C. L. Hubley-Kozey, "Muscle activation in therapeutic exercises to improve trunk stability," *Arch. Phys. Med. Rehabil.*, vol. 81, no. 10, pp. 1370–1379, 2000.

[65] W. Daniel, *Applied Nonparametric Statistics* (Duxbury Advanced Series in Statistics and Decision Sciences). Boston, MA, USA: PWS-KENT, 1990. [Online]. Available: https://books.google.it/books?id=ayDFAAAACAAJ

[66] B. N. Green *et al.*, "A scoping review of biopsychosocial risk factors and co-morbidities for common spinal disorders," *PLoS ONE*, vol. 13, no. 6, 2018, Art. no. e0197987.

[67] M. Capecci, L. Ciabattoni, F. Ferracuti, A. Monteriù, L. Romeo, and F. Verdini, "Collaborative design of a telerehabilitation system enabling virtual second opinion based on fuzzy logic," *IET Comput. Vis.*, vol. 12, pp. 502–512, Jun. 2018.

[68] A. Desjardins-Charbonneau, J.-S. Roy, J. Thibault, V. T. Ciccone, and F. Desmeules, "Acceptability of physiotherapists as primary care practitioners and advanced practice physiotherapists for care of patients with musculoskeletal disorders: A survey of a University community within the province of quebec," *BMC Musculoskeletal Disorders*, vol. 17, no. 1, p. 400, Sep. 2016.